

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 / Search time 13.7006 Seconds  
(Without alignments)  
222.664 Million cell updates/sec

Title: US-10-092-750-4  
Perfect score: 215  
Sequence: 1 TGEAIIIRLVALLBEAEV.....SDPALRSKLVRLSSDSFARL 46

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 6631800 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199	92.6	249	2	US-08-632-514C-11
2	199	92.6	249	3	US-09-188-177-11
3	66	30.7	1279	3	US-09-724-517-2
4	66	30.7	1279	4	US-09-641-807A-2
5	66	30.7	1279	4	US-09-723-086-2
6	58	27.0	13	2	US-08-632-514C-1
7	58	27.0	13	2	US-08-632-514C-5
8	58	27.0	13	2	US-08-632-514C-18
9	58	27.0	13	2	US-08-632-514C-24
10	58	27.0	13	3	US-09-188-177-1
11	58	27.0	13	3	US-09-188-177-5
12	58	27.0	13	3	US-09-188-177-18
13	58	27.0	13	3	US-09-188-177-24
14	57.5	26.7	452	2	US-08-686-599A-18
15	57.5	26.7	467	2	US-08-686-599A-17
16	57.5	26.7	493	2	US-08-686-599A-5
17	57.5	26.7	493	2	US-08-686-599A-16
18	56	26.0	149	4	US-09-488-039A-12503
19	56	26.0	736	4	US-09-583-110-4555
20	56	26.0	741	4	US-09-489-039A-8111
21	56	26.0	105	4	US-09-252-991A-23779
22	55	25.6	103	4	US-09-248-796A-18538
23	55	25.6	236	4	US-09-248-796A-15717
24	55	25.6	244	4	US-09-270-767-36375
25	55	25.6	244	4	US-09-270-767-36375
26	55	25.6	373	4	US-09-252-991A-30628
27	54.5	25.3	248	4	US-09-328-352-6994

28	54	25.1	13	2	US-08-632-514C-2	Sequence 2, Appl
29	54	25.1	13	2	US-08-632-514C-6	Sequence 6, Appl
30	54	25.1	13	2	US-08-632-514C-26	Sequence 26, Appl
31	54	25.1	13	3	US-09-188-177-2	Sequence 2, Appl
32	54	25.1	13	3	US-09-188-177-6	Sequence 6, Appl
33	54	25.1	13	3	US-09-188-177-26	Sequence 26, Appl
34	54	25.1	13	3	US-09-328-352-6547	Sequence 6547, Ap
35	54	25.1	1397	4	US-09-248-796A-21477	Sequence 21477, A
36	54	25.1	1274	4	US-09-095-443-2	Sequence 2, Appl
37	53.5	24.9	915	4	US-09-252-991A-22344	Sequence 22344, A
38	53	24.7	13	2	US-08-632-514C-3	Sequence 3, Appl
39	53	24.7	13	2	US-08-632-514C-7	Sequence 7, Appl
40	53	24.7	13	2	US-08-632-514C-27	Sequence 27, Appl
41	53	24.7	13	3	US-09-188-177-3	Sequence 3, Appl
42	53	24.7	13	3	US-09-188-177-7	Sequence 7, Appl
43	53	24.7	13	3	US-09-188-177-27	Sequence 27, Appl
44	53	24.7	600	3	US-09-352-159-36	Sequence 36, Appl
45	53	24.7	600	3	US-09-352-159-38	Sequence 38, Appl

## ALIGNMENTS

```
RESULT 1
US-08-632-514C-11
Sequence 11, Application US/08632514C
Patent No. 5834234
GENERAL INFORMATION:
APPLICANT: GALLO, Gregory J.
TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BK
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/632,514C
FILING DATE: 29-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, Henry N.
REGISTRATION NUMBER: 32,073
REFERENCE/DOCKET NUMBER: 104322.188
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8459
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-632-514C-11
Query Match 92.6%; Score 199; DB 2; Length 249;
Best Local Similarity 95.7%; Pred. No. 1.9e-20;
Matches 44; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Sequence 11, Application US/09188177  
Patent No. 6057132  
GENERAL INFORMATION:  
APPLICANT: GALLO, Gregory J.  
TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BAX  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington, D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/188,177  
FILING DATE: 2  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, Henry N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.188  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)942-8459  
TELEFAX: (202)942-8484  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 249 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-188-177-11

Query Match 92.6%; Score 199; DB 3; Length 249;  
Best Local Similarity 95.7%; Pred. No. 1.9e-20;  
Matches 44; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TKEALIRLVALLIEEAEVINOGLASDPALRSKLVRLSSPRL 46  
DB 119 TEKALIRLVALLIEEAEVINOGLASDPALRSKLVRLSSPRL 164

RESULT 3  
US-09-724-517-2  
Sequence 2, Application US/09724517  
Patent No. 6379941  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
TITLE OF INVENTION: Freedman, Richard  
TITLE OF INVENTION: No. 6379941e1 motor proteins and methods for  
FILE REFERENCE: 1031  
CURRENT APPLICATION NUMBER: US/09/724,517  
CURRENT FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: US/09/641,807  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1279  
TYPE: PRT  
ORGANISM: Human  
FEATURE:  
NAME/KEY: VARIANT (436)  
LOCATION: (409)  
OTHER INFORMATION: Xaa = any amino acid  
US-09-724-517-2

Query Match 30.7%; Score 66; DB 3; Length 1279;

Best Local Similarity 36.8%; Pred. No. 1.4;  
Matches 14; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 3 KEALIRLVALLIEEAEVINOGLASDPALRSKLVRLSS 40  
DB 830 REAIVSKKALIQEKSHLENKLRSSQALNTDSLKIST 867

RESULT 4  
US-09-641-807A-2  
Sequence 2, Application US/09641807A  
Patent No. 6440731  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
TITLE OF INVENTION: Freedman, Richard  
TITLE OF INVENTION: No. 6440731e1 motor proteins and methods for  
FILE REFERENCE: 1031  
CURRENT APPLICATION NUMBER: US/09/641,807A  
CURRENT FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1279  
TYPE: PRT  
ORGANISM: Human  
FEATURE:  
NAME/KEY: VARIANT (446)  
LOCATION: (409)  
OTHER INFORMATION: Xaa = any amino acid  
US-09-641-807A-2

Query Match 30.7%; Score 66; DB 4; Length 1279;  
Best Local Similarity 36.8%; Pred. No. 1.4;  
Matches 14; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 3 KEALIRLVALLIEEAEVINOGLASDPALRSKLVRLSS 40  
DB 830 REAIVSKKALIQEKSHLENKLRSSQALNTDSLKIST 867

RESULT 5  
US-09-723-096-2  
Sequence 2, Application US/09723096  
Patent No. 6448026  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
TITLE OF INVENTION: Freedman, Richard  
TITLE OF INVENTION: No. 6448026e1 motor proteins and methods for  
FILE REFERENCE: 1031  
CURRENT APPLICATION NUMBER: US/09/723,096  
CURRENT FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: US/09/641,807  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1279  
TYPE: PRT  
ORGANISM: Human  
FEATURE:  
NAME/KEY: VARIANT (436)  
LOCATION: (409)  
OTHER INFORMATION: Xaa = any amino acid  
US-09-723-096-2

Query Match 30.7%; Score 66; DB 4; Length 1279;  
Best Local Similarity 36.8%; Pred. No. 1.4;  
Matches 14; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 3 KEALIRLVALLIEEAEVINOGLASDPALRSKLVRLSS 40  
DB 830 REAIVSKKALIQEKSHLENKLRSSQALNTDSLKIST 867

Db 830 REAIVSKKALLQEKSHENKKLRSSQALNTDSLKIST 867

RESULT 6  
US-08-632-514C-1  
Sequence 1, Application US/08632514C  
Patent No. 5834234  
GENERAL INFORMATION:  
APPLICANT: GALLO, Gregory J.  
TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington, D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/632,514C  
FILING DATE: 29-MAY-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, Henry N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.188  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)942-8459  
TELEFAX: (202)942-8484  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acid  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-632-514C-1

Query Match 27.0%; Score 58; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.062; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRVLALLEEAE 19  
Db 1 LRRVLALLEEAE 13

RESULT 7  
US-08-632-514C-5  
Sequence 5, Application US/08632514C  
Patent No. 5834234  
GENERAL INFORMATION:  
APPLICANT: GALLO, Gregory J.  
TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington, D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/632,514C  
FILING DATE: 29-MAY-1996

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, Henry N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.188  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)942-8459  
TELEFAX: (202)942-8484  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-632-514C-5

Query Match 27.0%; Score 58; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.062; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRVLALLEEAE 19  
Db 1 LRRVLALLEEAE 13

RESULT 8  
US-08-632-514C-18  
Sequence 18, Application US/08632514C  
Patent No. 5834234  
GENERAL INFORMATION:  
APPLICANT: GALLO, Gregory J.  
TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington, D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/632,514C  
FILING DATE: 29-MAY-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, Henry N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.188  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)942-8459  
TELEFAX: (202)942-8484  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-632-514C-18

Query Match 27.0%; Score 58; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.062; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRVLALLEEAE 19  
Db 1 LRRVLALLEEAE 13

## RESULT 9

US-08-632-514C-24

; Sequence 24, Application US/08632514C  
; Patent No. 5834234

; GENERAL INFORMATION:

; APPLICANT: GALLO, Gregory J.

; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hale and Dorr

; STREET: 1455 Pennsylvania Avenue, N.W.

; CITY: Washington, D.C.

; COUNTRY: U.S.A.

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Releasee #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/632,514C

; FILING DATE: 29-MAY-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: WIXON, Henry N.

; REGISTRATION NUMBER: 32,073

; REFERENCE/DOCKET NUMBER: 104322.188

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)942-8459

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 13 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-632-514C-24

QY 7 LRRVLALLEEAE 19

Db 1 LRRVLALLEEAE 13

Query Match 27.0%; Score 58; DB 2; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.062;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10

US-09-188-177-1

; Sequence 1, Application US/09188177

; Patent No. 6057132

; GENERAL INFORMATION:

; APPLICANT: GALLO, Gregory J.

; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hale and Dorr

; STREET: 1455 Pennsylvania Avenue, N.W.

; CITY: Washington, D.C.

; COUNTRY: U.S.A.

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Releasee #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/188,177

; FILING DATE: 2

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

NAME: WIXON, Henry N.

REGISTRATION NUMBER: 32,073

REFERENCE/DOCKET NUMBER: 104322.188

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)942-8459

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acid

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-188-177-1

QY 7 LRRVLALLEEAE 19

Db 1 LRRVLALLEEAE 13

Query Match 27.0%; Score 58; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.062;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11

US-09-188-177-5

; Sequence 5, Application US/09188177

; Patent No. 6057132

; GENERAL INFORMATION:

; APPLICANT: GALLO, Gregory J.

; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hale and Dorr

; STREET: 1455 Pennsylvania Avenue, N.W.

; CITY: Washington, D.C.

; COUNTRY: U.S.A.

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Releasee #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/188,177

; FILING DATE: 2

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: WIXON, Henry N.

; REGISTRATION NUMBER: 32,073

; REFERENCE/DOCKET NUMBER: 104322.188

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)942-8459

; INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-188-177-5

QY 7 LRRVLALLEEAE 19

Db 1 LRRVLALLEEAE 13

Query Match 27.0%; Score 58; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.062;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12



US-09-188-177-18  
Sequence 18, Application US/09188177  
Patent No. 6057132  
GENERAL INFORMATION:  
APPLICANT: GALLO, Gregory J.  
TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington, D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/188,177  
FILING DATE: 2  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, Henry N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.188  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)942-8459  
TELEFAX: (202)942-8484  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-188-177-18

Query Match 27.0%; Score 58; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred.No. 0.062; 0; Indels 0; Gaps 0;  
Matches 13; Conservative 0; Mismatches 0;

QY 7 LRLVALLEBEAE 19  
DB 1 LRLVALLEBEAE 13

RESULT 13  
US-09-188-177-24  
Sequence 24, Application US/09188177  
Patent No. 6057132  
GENERAL INFORMATION:  
APPLICANT: GALLO, Gregory J.  
TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington, D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/188,177  
FILING DATE: 2  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, Henry N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.188  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)942-8459  
TELEFAX: (202)942-8484  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-188-177-18

REFERENCE/DOCKET NUMBER: 104322.188  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)942-8459  
TELEFAX: (202)942-8484  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-188-177-24

Query Match 27.0%; Score 58; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred.No. 0.062; 0; Indels 0; Gaps 0;  
Matches 13; Conservative 0; Mismatches 0;

QY 7 LRLVALLEBEAE 19  
DB 1 LRLVALLEBEAE 13

RESULT 14  
US-08-686-599A-18  
Sequence 18, Application US/08686599A  
Patent No. 5891439  
GENERAL INFORMATION:  
APPLICANT: Olsson, Thomas  
APPLICANT: Valdivia, Tushar  
APPLICANT: Bakhet, Abdel-Molz  
APPLICANT: Kristensson, Kristofer  
APPLICANT: Donelson, John E.  
TITLE OF INVENTION: LYMPHOCYTE STIMULATING FACTOR  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314-3187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/686,599A  
FILING DATE: 26-JUL-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Duffett, Benton S.  
REGISTRATION NUMBER: 22,030  
REFERENCE/DOCKET NUMBER: 003300-383  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: TLIR  
US-08-686-599A-18

Query Match 26.7%; Score 57.5; DB 2; Length 452;  
Best Local Similarity 48.4%; Pred.No. 6.3; 10; Indels 1; Gaps 1;  
Matches 15; Conservative 5; Mismatches 374

QY 2 GK-BAIRLRLVALLEBEAEVINOKLASDPAL 31  
DB 344 GKFSALRQAVVVERNEVLEQXILIESHAL 374

## RESULT 15

US-08-686-599A-17  
 ; Sequence 17, Application US/08686599A  
 ; Patent No. 5891439  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Olsson, Thomas  
 ; APPLICANT: Vaidya, Tushar  
 ; APPLICANT: Bakhtel, Abdel-Molz  
 ; APPLICANT: Kristenson, Krista  
 ; APPLICANT: Donelson, John E.  
 ; TITLE OF INVENTION: LYMPHOCYTE STIMULATING FACTOR  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP  
 ; STREET: 699 Prince Street  
 ; CITY: Alexandria  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22314-3187  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/686,599A  
 ; FILING DATE: 26-JUL-1996  
 ; CLASSIFICATION: 536  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Duffett, Benton S.  
 ; REGISTRATION NUMBER: 22,030  
 ; REFERENCE/DOCKET NUMBER: 003300-383  
 ; INFORMATION FOR SEQ ID NO: 17:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 467 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: T/TF  
 ; US-08-686-599A-17

Query Match 26.7%; Score 57.5; DB 2; Length 467;  
 Best Local Similarity 48.4%; Fred. No. 6.6;  
 Matches 15; Conservative 5; Mismatches 10; Indels 1; Gaps 1;  
 QY 2 GK-EALIRRLVALLLEEAEVYNQKLASDPAL 31  
 Db 344 GKFSALRQAVWVVEERNEVLTQCKLIESHAL 374

Search completed: November 10, 2004, 12:32:19  
 Job time: 13.7006 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 / Search time 44.3673 Seconds  
(without alignments)  
366.225 Million cell updates/sec

Title: US-10-092-750-4  
Perfect score: 215  
Sequence: 1 TGEKALIRLVALLLEENAEV.....SDPALRSKLVRLSSDFAHLL 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
6:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
7:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
8:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
17:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
18:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
19:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
20:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215	100.0	46	US-10-092-750-4	Sequence 4, Appl1
2	199	92.6	334	US-09-953-342-24	Sequence 24, Appl1
3	142	66.0	237	US-09-925-301-867	Sequence 867, App
4	104.5	48.6	159	US-09-890-688-110	Sequence 110, App
5	67	31.2	830	US-09-866-034-7	Sequence 7, Appl1
6	67	31.2	830	US-10-033-246-7	Sequence 7, Appl1
7	67	31.2	830	US-10-033-301-7	Sequence 7, Appl1
8	67	31.2	830	US-10-033-326-7	Sequence 7, Appl1
9	67	31.2	830	US-10-033-245-7	Sequence 7, Appl1
10	67	31.2	830	US-10-033-223-7	Sequence 7, Appl1
11	67	31.2	830	US-10-033-167-7	Sequence 7, Appl1
12	67	31.2	830	US-10-033-244-7	Sequence 7, Appl1
13	67	31.2	830	US-10-033-435-7	Sequence 7, Appl1

ALIGNMENTS

14	67	31.2	830	14	US-10-032-990-7	Sequence 7, Appl1
15	67	31.2	830	14	US-10-032-996-7	Sequence 7, Appl1
16	67	31.2	830	14	US-10-033-996-7	Sequence 7, Appl1
17	67	31.2	830	14	US-10-210-951-66	Sequence 66, Appl1
18	67	31.2	830	14	US-10-211-884-66	Sequence 66, Appl1
19	67	31.2	830	14	US-10-211-858-66	Sequence 66, Appl1
20	67	31.2	830	14	US-10-439-249-7	Sequence 7, Appl1
21	67	31.2	830	16	US-10-692-072-7	Sequence 7, Appl1
22	67	31.2	1324	15	US-10-287-226-314	Sequence 314, App
23	66	30.7	1401	15	US-10-287-226-142	Sequence 142, App
24	61.5	28.6	357	14	US-10-369-493-2652	Sequence 2652, Ap
25	60	27.9	643	15	US-10-382-122A-54450	Sequence 54450, A
26	58.5	27.2	338	14	US-10-369-493-18919	Sequence 18919, A
27	58.5	27.2	353	14	US-10-369-493-19950	Sequence 19950, A
28	57.5	26.7	575	15	US-10-282-122A-62303	Sequence 62303, A
29	57.5	26.7	575	15	US-10-282-122A-64407	Sequence 64407, A
30	57.5	26.7	933	17	US-10-425-115-350329	Sequence 350329, A
31	57	26.5	321	14	US-10-369-493-3708	Sequence 3708, Ap
32	57	26.5	357	15	US-10-282-122A-50703	Sequence 50703, A
33	57	26.5	709	16	US-10-437-963-109502	Sequence 109502, A
34	57	26.5	2045	15	US-10-282-122A-74463	Sequence 74463, A
35	56.5	26.3	269	16	US-10-437-963-192646	Sequence 192646, A
36	56.5	26.3	334	16	US-10-437-963-192648	Sequence 192648, A
37	56	26.0	348	14	US-10-369-493-14599	Sequence 14599, A
38	56	26.0	376	14	US-10-369-493-11645	Sequence 11645, A
39	56	26.0	442	15	US-10-188-186-124	Sequence 126, App
40	56	26.0	442	15	US-10-188-186-126	Sequence 126, App
41	56	26.0	480	15	US-10-408-456-30	Sequence 30, Appl
42	56	26.0	480	15	US-10-408-456-32	Sequence 32, Appl
43	56	26.0	863	15	US-10-282-122A-58589	Sequence 58589, A
44	56	26.0	1179	9	US-09-815-242-13262	Sequence 13262, A
45	56	26.0	1179	9	US-09-815-242-13608	Sequence 13608, A

RESULT 1  
US-10-092-750-4  
; Sequence 4, Application US/10092750  
; Publication No. US2003032157A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; APPLICANT: Alpin, Naita  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL  
; FILE REFERENCE: 50036/050002  
; CURRENT APPLICATION NUMBER: US/10/092,750  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: US 60/274,526  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 46  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-750-4

Query Match 100.0%; Score 215; DB 14; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.7e-21;  
Matches 46; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;

QY 1 TGEKALIRLVALLLEENAEVINOGLASDPALRSKLVRLSSDFAHLL 46  
DB 1 TGEKALIRLVALLLEENAEVINOGLASDPALRSKLVRLSSDFAHLL 46

RESULT 2  
US-09-953-342-24  
; Sequence 24, Application US/09953342  
; Patent No. US20020106735A1  
; GENERAL INFORMATION:

```

; APPLICANT: Scorilas, Andreas
; APPLICANT: Diamandis, Eleftherios
; TITLE OF INVENTION: NOVEL BCL-2 RELATED PROLINE RICH PROTEIN (BPR)
; FILE REFERENCE: 11757.52USU1
; CURRENT APPLICATION NUMBER: US/09/953,342
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/233,026
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-953-342-24

```

```

Query Match          92.6%; Score 199; DB 9; Length 334;
Best Local Similarity 95.7%; Pred. No. 2.8e-18;
Matches 44; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 TGKATLRRLVALLBEAEVYNOKLASDPALRSKYVRLSSDPFAH 46
Db      204 TEKATLRRLVALLBEAEVYNOKLASDPALRSKYVRLSSDPFAH 249

```

```

RESULT 3
US-09-925-301-867
; Sequence 867, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 867
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-867

```

```

Query Match          66.0%; Score 142; DB 9; Length 237;
Best Local Similarity 93.9%; Pred. No. 9e-11;
Matches 31; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 TGKATLRRLVALLBEAEVYNOKLASDPALRS 33
Db      160 TEKATLRRLVALLBEAEVYNOKLASDPALRT 192

```

```

RESULT 4
US-09-890-688-110
; Sequence 110, Application US/09890688
; Publication No. US20030144475A1
; GENERAL INFORMATION:
; APPLICANT: Seishi KATO
; APPLICANT: Chikashi EGUCHI
; APPLICANT: Mahoro SAKI
; TITLE OF INVENTION: Human Proteins and cDNAs thereof
; FILE REFERENCE: 2001-1102A/MMC/00653
; CURRENT APPLICATION NUMBER: US/09/890,688
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: JP 11-346863
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 11-34684
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 2000-31062

```

```

; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: JP 2000-34091
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-34090
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-35829
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-35899
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-71161
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: JP 2000-160851
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 110
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-890-688-110

```

```

Query Match          48.6%; Score 104.5; DB 10; Length 159;
Best Local Similarity 81.2%; Pred. No. 6.2e-06;
Matches 26; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

```

```

QY      1 TGKATLRRLVALLBEAEVYNOK--LASDP 29
Db      120 TEKATLRRLVALLBEAEVYNOKEGITAVSP 151

```

```

RESULT 5
US-09-866-034-7
; Sequence 7, Application US/09866034
; Publication No. US20030170864A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2930R1C1
; CURRENT APPLICATION NUMBER: US/09/866,034
; CURRENT FILING DATE: 2001-05-25
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 7
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-034-7

```

```

Query Match          31.2%; Score 67; DB 10; Length 830;
Best Local Similarity 39.5%; Pred. No. 5.6;
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```

```

QY      3 KEATLRRLVALLBEAEVYNOKLASDPALRSKYVRLSS 40
Db      426 REATLAKKALMCKETGELSKRLSSQALNEDIVRSS 463

```

```

RESULT 6
US-10-033-246-7
; Sequence 7, Application US/10033246

```

Publication No. US20020098505A1  
 GENERAL INFORMATION:  
 APPLICANT: Botstein, David  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Ferrara, Napoleone  
 APPLICANT: Fong, Sherman  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Pan, James  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Wood, William I.  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 TITLE OF INVENTION: Acids Encoding the Same  
 FILE REFERENCE: P2930R1C12  
 CURRENT APPLICATION NUMBER: US/10/033,246  
 CURRENT FILING DATE: 2001-12-28  
 PRIOR APPLICATION NUMBER: 60/095,325  
 PRIOR FILING DATE: 1998-08-04  
 PRIOR APPLICATION NUMBER: 60/112,851  
 PRIOR FILING DATE: 1998-12-16  
 PRIOR APPLICATION NUMBER: 60/113,145  
 PRIOR FILING DATE: 1998-12-16  
 PRIOR APPLICATION NUMBER: 60/113,511  
 PRIOR FILING DATE: 1998-12-22  
 PRIOR APPLICATION NUMBER: 60/115,558  
 PRIOR FILING DATE: 1999-01-12  
 PRIOR APPLICATION NUMBER: 60/115,565  
 PRIOR FILING DATE: 1999-01-12  
 PRIOR APPLICATION NUMBER: 60/115,733  
 PRIOR FILING DATE: 1999-01-12  
 PRIOR APPLICATION NUMBER: 60/119,341  
 PRIOR FILING DATE: 1999-02-09  
 PRIOR APPLICATION NUMBER: 60/119,537  
 PRIOR FILING DATE: 1999-02-10  
 PRIOR APPLICATION NUMBER: 60/119,965  
 PRIOR FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: 60/162,506  
 PRIOR FILING DATE: 1999-10-29  
 PRIOR APPLICATION NUMBER: 60/170,262  
 PRIOR FILING DATE: 1999-12-09  
 PRIOR APPLICATION NUMBER: 60/187,202  
 PRIOR FILING DATE: 2000-03-03  
 PRIOR APPLICATION NUMBER: PCT/US99/12252  
 PRIOR FILING DATE: 1999-06-02  
 PRIOR APPLICATION NUMBER: PCT/US99/28634  
 PRIOR FILING DATE: 1999-12-01  
 PRIOR APPLICATION NUMBER: PCT/US99/28551  
 PRIOR FILING DATE: 1999-12-02  
 PRIOR APPLICATION NUMBER: PCT/US00/03565  
 PRIOR FILING DATE: 2000-02-11  
 PRIOR APPLICATION NUMBER: PCT/US00/04414  
 PRIOR FILING DATE: 2000-02-22  
 PRIOR APPLICATION NUMBER: PCT/US00/05841  
 PRIOR FILING DATE: 2000-03-02  
 PRIOR APPLICATION NUMBER: PCT/US00/06439  
 PRIOR FILING DATE: 2000-03-30  
 PRIOR APPLICATION NUMBER: PCT/US00/14941  
 PRIOR FILING DATE: 2000-05-30  
 PRIOR APPLICATION NUMBER: PCT/US00/15264  
 PRIOR FILING DATE: 2000-06-02  
 PRIOR APPLICATION NUMBER: PCT/US00/32678  
 PRIOR FILING DATE: 2000-12-01  
 NUMBER OF SEQ ID NOS: 38  
 SEQ ID NO: 7  
 LENGTH: 830  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-033-246-7

Query Match 31.2%; Score 67; DB 13; Length 830;  
 Best Local Similarity 39.5%; Pred. No. 5.6;  
 Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;  
 QY 3 KEAIRRLVALLREAEAVINOKLASDPALRSKYRLSS 40  
 Db 426 REAILARKEALMOKSKTGSLRSQALNEDIVRVS 463

RESULT 7  
 US-10-033-301-7  
 Sequence 7, Application US/10033301  
 Publication No. US20020098506A1  
 GENERAL INFORMATION:  
 APPLICANT: Botstein, David  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Ferrara, Napoleone  
 APPLICANT: Fong, Sherman  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Pan, James  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Wood, William I.  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 TITLE OF INVENTION: Acids Encoding the Same  
 FILE REFERENCE: P2930R1C6  
 CURRENT APPLICATION NUMBER: US/10/033,301  
 CURRENT FILING DATE: 2001-12-27  
 PRIOR APPLICATION NUMBER: 60/095,325  
 PRIOR FILING DATE: 1998-08-04  
 PRIOR APPLICATION NUMBER: 60/112,851  
 PRIOR FILING DATE: 1998-12-16  
 PRIOR APPLICATION NUMBER: 60/113,145  
 PRIOR FILING DATE: 1998-12-16  
 PRIOR APPLICATION NUMBER: 60/113,511  
 PRIOR FILING DATE: 1998-12-22  
 PRIOR APPLICATION NUMBER: 60/115,558  
 PRIOR FILING DATE: 1999-01-12  
 PRIOR APPLICATION NUMBER: 60/115,565  
 PRIOR FILING DATE: 1999-01-12  
 PRIOR APPLICATION NUMBER: 60/115,733  
 PRIOR FILING DATE: 1999-01-12  
 PRIOR APPLICATION NUMBER: 60/119,341  
 PRIOR FILING DATE: 1999-02-09  
 PRIOR APPLICATION NUMBER: 60/119,537  
 PRIOR FILING DATE: 1999-02-10  
 PRIOR APPLICATION NUMBER: 60/119,965  
 PRIOR FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: 60/162,506  
 PRIOR FILING DATE: 1999-10-29  
 PRIOR APPLICATION NUMBER: 60/170,262  
 PRIOR FILING DATE: 1999-12-09  
 PRIOR APPLICATION NUMBER: 60/187,202  
 PRIOR FILING DATE: 2000-03-03  
 PRIOR APPLICATION NUMBER: PCT/US99/12252  
 PRIOR FILING DATE: 1999-06-02  
 PRIOR APPLICATION NUMBER: PCT/US99/28634  
 PRIOR FILING DATE: 1999-12-01  
 PRIOR APPLICATION NUMBER: PCT/US99/28551  
 PRIOR FILING DATE: 1999-12-02  
 PRIOR APPLICATION NUMBER: PCT/US00/03565  
 PRIOR FILING DATE: 2000-02-11  
 PRIOR APPLICATION NUMBER: PCT/US00/04414  
 PRIOR FILING DATE: 2000-02-22  
 PRIOR APPLICATION NUMBER: PCT/US00/05841  
 PRIOR FILING DATE: 2000-03-02  
 PRIOR APPLICATION NUMBER: PCT/US00/08439  
 PRIOR FILING DATE: 2000-03-30  
 PRIOR APPLICATION NUMBER: PCT/US00/14941

```

; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 7
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-301-7

Query Match      31.2%; Score 67; DB 13; Length 830;
Best Local Similarity 39.5%; Pred. No. 5.6;
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Cy      3 KEAIRRLVALLKEEFAVINOKLASDPALRSKLVRLSS 40
Db      426 REAILAKKELMOEKTGLSKRLRSSQALNEDIVRVSS 463

RESULT 8
US-10-033-326-7
; Sequence 7, Application US/10033326
; Publication No. US20020098507A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2930R1C8
; CURRENT FILING DATE: 2001-12-27
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
```

```

; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 7
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-326-7

Query Match      31.2%; Score 67; DB 13; Length 830;
Best Local Similarity 39.5%; Pred. No. 5.6;
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Cy      3 KEAIRRLVALLKEEFAVINOKLASDPALRSKLVRLSS 40
Db      426 REAILAKKELMOEKTGLSKRLRSSQALNEDIVRVSS 463

RESULT 9
US-10-033-245-7
; Sequence 7, Application US/10033245
; Publication No. US20020160392A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2930R1C7
; CURRENT FILING DATE: 2001-12-27
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,537
```

```

; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 7
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-245-7

```

```

Query Match          31.2%; Score 67; DB 13; Length 830;
Best Local Similarity 39.5%; Pred. No. 5.6;
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```

```

; RESULT 10
; US-10-033-223-7
; Sequence 7, Application US/10033223
; Publication No. US20020164646A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2930R1C9
; CURRENT APPLICATION NUMBER: US/10/033,223
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145

```

```

; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 7
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-223-7

```

```

Query Match          31.2%; Score 67; DB 13; Length 830;
Best Local Similarity 39.5%; Pred. No. 5.6;
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```

```

; RESULT 11
; US-10-033-167-7
; Sequence 7, Application US/10033167
; Publication No. US20020182618A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel

```

```

; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2930R1C10
; CURRENT APPLICATION NUMBER: US/10/033,167
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 7
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-167-7

Query Match          31.2%; Score 67; DB 13; Length 830;
Best Local Similarity 39.5%; Pred. No. 5.6;
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```

```

; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gunney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2930R1C2
; CURRENT APPLICATION NUMBER: US/10/033,244
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 7
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-244-7

Query Match          31.2%; Score 67; DB 13; Length 830;

```



Best Local Similarity 39.5%; Pred. No. 5.6;  
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;  
Qy 3 KEAIIKRLVALLKEEAEVINGKLAADPALRSKLVRLSS 40  
Db 426 REAIIAKKELMKEKTEGHSKELRSQALNEDIYRVSS 463

RESULT 13  
US-10-033-435-7  
Sequence 7, Application US/10033435  
Publication No. US20030027256A1  
GENERAL INFORMATION:  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2930R1C5  
CURRENT APPLICATION NUMBER: US/10/033,435  
PRIOR FILING DATE: 2001-12-27  
PRIOR APPLICATION NUMBER: 60/095,325  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/112,851  
PRIOR FILING DATE: 1998-12-16  
PRIOR APPLICATION NUMBER: 60/113,145  
PRIOR FILING DATE: 1998-12-16  
PRIOR APPLICATION NUMBER: 60/113,511  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/115,558  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115,565  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115,733  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/119,341  
PRIOR FILING DATE: 1999-02-09  
PRIOR APPLICATION NUMBER: 60/119,537  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/119,965  
PRIOR FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: 60/162,506  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 60/170,262  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 60/187,202  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: PCT/US99/12252  
PRIOR FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: PCT/US99/28634  
PRIOR FILING DATE: 1999-12-01  
PRIOR APPLICATION NUMBER: PCT/US99/28551  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US00/03565  
PRIOR FILING DATE: 2000-02-11  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: PCT/US00/05841  
PRIOR FILING DATE: 2000-03-02  
PRIOR APPLICATION NUMBER: PCT/US00/08439  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: PCT/US00/14941  
PRIOR FILING DATE: 2000-05-30

PRIOR APPLICATION NUMBER: PCT/US00/15264  
PRIOR FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: PCT/US00/32678  
PRIOR FILING DATE: 2000-12-01  
NUMBER OF SEQ ID NOS: 38  
SEQ ID NO 7  
LENGTH: 830  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-033-435-7

Query Match 31.2%; Score 67; DB 14; Length 830;  
Best Local Similarity 39.5%; Pred. No. 5.6;  
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;  
Qy 3 KEAIIKRLVALLKEEAEVINGKLAADPALRSKLVRLSS 40  
Db 426 REAIIAKKELMKEKTEGHSKELRSQALNEDIYRVSS 463

RESULT 14  
US-10-032-990-7  
Sequence 7, Application US/10032990  
Publication No. US20030032060A1  
GENERAL INFORMATION:  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2930R1C11  
CURRENT APPLICATION NUMBER: US/10/032,990  
PRIOR FILING DATE: 2001-12-27  
PRIOR APPLICATION NUMBER: 60/095,325  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/112,851  
PRIOR FILING DATE: 1998-12-16  
PRIOR APPLICATION NUMBER: 60/113,145  
PRIOR FILING DATE: 1998-12-16  
PRIOR APPLICATION NUMBER: 60/113,511  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/115,558  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115,565  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115,733  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/119,341  
PRIOR FILING DATE: 1999-02-09  
PRIOR APPLICATION NUMBER: 60/119,537  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/119,965  
PRIOR FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: 60/162,506  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 60/170,262  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 60/187,202  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: PCT/US99/12252  
PRIOR FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: PCT/US99/28634  
PRIOR FILING DATE: 1999-12-01

```

1 CURRENT APPLICATION NUMBER: 10/033,996
2 CURRENT FILING DATE: 2001-12-27
3 PRIOR APPLICATION NUMBER: 60/095,325
4 PRIOR FILING DATE: 1998-08-04
5 PRIOR APPLICATION NUMBER: 60/112,851
6 PRIOR FILING DATE: 1998-12-16
7 PRIOR APPLICATION NUMBER: 60/113,145
8 PRIOR FILING DATE: 1998-12-16
9 PRIOR APPLICATION NUMBER: 60/113,511
10 PRIOR FILING DATE: 1998-12-22
11 PRIOR APPLICATION NUMBER: 60/115,558
12 PRIOR FILING DATE: 1999-01-12
13 PRIOR APPLICATION NUMBER: 60/115,565
14 PRIOR FILING DATE: 1999-01-12
15 PRIOR APPLICATION NUMBER: 60/115,733
16 PRIOR FILING DATE: 1999-01-12
17 PRIOR APPLICATION NUMBER: 60/119,341
18 PRIOR FILING DATE: 1999-02-09
19 PRIOR APPLICATION NUMBER: 60/119,537
20 PRIOR FILING DATE: 1999-02-10

```

Search completed: November 10, 2004, 16:35:39  
Job time : 45.3673 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 10, 2004, 11:36:51 ; Search time 9.29938 Seconds  
(without alignments)  
475.942 Million cell updates/sec

Title: US-10-092-750-4  
Perfect score: 215  
Sequence: 1 TGEKALRLVALLEBEAFV.....SDPALRSKLVRLSSDSFAHL 46  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	30.7	833	2	T43446 hypothetical prote
2	64	29.8	2672	2	A48126 translation activa
3	61.5	28.6	357	2	T46868 protoporphyrin IX
4	61.5	28.6	357	2	S75252 protoporphyrin IX
5	61	28.4	2581	2	AF2545 hypothetical prote
6	60	27.9	643	2	B81361 ABC transport syst
7	59	27.4	356	2	B82880 probable transcrip
8	58.5	27.2	374	2	AH1825 protoporphyrin IX
9	58	27.0	171	2	E75562 Mult/ndix family
10	57.5	26.7	470	2	D63994 phosphoribosylamin
11	57.5	26.7	575	2	B70612 probable recd prot
12	57	26.5	261	2	S52110 septaplexin reduct
13	57	26.5	504	2	AC1740 S. pyogenes RefA x
14	56	26.0	259	2	E72224 ABC transporter, A
15	56	26.0	412	2	AF2979 acyl-CoA dehydroge
16	56	26.0	451	2	G98303 hypothetical prote
17	56	26.0	480	1	DCHUA aromatic-L-amino-a
18	56	26.0	863	2	B64138 uridylyltransferas
19	56	26.0	1004	2	H87112 glutamate-ammonia-
20	56	26.0	1179	2	G95144 conserved hypthe
21	56	26.0	1179	2	E98012 hypothetical prote
22	56	26.0	1802	2	H88444 protein C2E6.12 l
23	56	26.0	2149	2	T47655 hypothetical prote
24	55.5	25.8	840	2	I48968 major fibrous shea
25	55.5	25.8	849	2	I49060 major 80,000 Mr fi
26	55.5	25.8	1877	2	T40550 probable phosphati
27	55.5	25.8	3562	2	A47171 chondroitin sulfat
28	55	25.6	138	2	AB0131 conserved hypthe
29	55	25.6	140	2	AC0532 probable release f

30	55	25.6	298	2	D57444 chemotaxis MotB pr
31	55	25.6	332	2	H31069 probable c4-dicarb
32	55	25.6	416	2	AB2911 peptidoglycan acet
33	55	25.6	440	2	H97685 hypothetical prote
34	55	25.6	480	1	DCRTA aromatic-L-amino-a
35	55	25.6	608	2	S09790 hypothetical prote
36	55	25.6	625	2	B42361 DNA-directed RNA p
37	55	25.6	625	2	AE2005 RNA polymerase gam
38	55	25.6	825	2	D6151 F22M8.8 protein -
39	55	25.3	275	2	A99517 ABC transporter at
40	54.5	25.3	367	2	A33950 yopM protein - Y
41	54.5	25.3	409	2	T43599 yop targeted effec
42	54.5	25.3	483	2	D97535 nitrogen regulator
43	54.5	25.3	483	2	AF2754 two component resp
44	54.5	25.3	483	2	I19719 nitrogen regulator
45	54.5	25.3	642	2	T42133 type II secretion

## ALIGNMENTS

## RESULT 1

T43446  
hypothetical protein DKFZp434D0917.1 - human (fragments)  
C/Species: Homo sapiens (man)  
C/Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
R/Blocher, H.; Bocker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, December 1999  
A/Reference number: Z22513  
A/Accession: T43446  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-167;168-833 <AAA>  
A/Cross-references: UNIPROT:Q9UF54; EMBL:AL133654  
A/Experimental source: adult testis; clone DKFZp434D0917  
A/Note: the cDNA sequence contains a -1 frameshift near codon 167  
C/Genetics:  
A/Note: DKFZp434D0917.1

Query Match 30.7%; Score 66; DB 2; Length 833;  
Best Local Similarity 36.8%; Pred. No. 5.9;  
Matches 14; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 3 KEAIRRLVALLEBEAFVYNOKLASDPALRSKLVRLSS 40  
DB 384 REAIVSKKVALLOKSHLENKRSQALNTDSIKIST 421

## RESULT 2

T48126  
translation activator GCN1 - yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein G1318; protein YG1195w  
C/Species: Saccharomyces cerevisiae  
C/Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
A/Accession: A48126; S62050; S64212  
R/Marton, M.J.; Crouch, D.; Himebusch, A.G.  
Mol. Cell. Biol. 13, 3541-3556, 1993  
A/Title: GCN1, a translational activator of GCN4 in Saccharomyces cerevisiae, is requi  
A/Reference number: A48126; WUID:93268304; PMID:8497269  
A/Accession: A48126  
A/Molecule type: DNA  
A/Residues: 1-2672 <NAR>  
A/Cross-references: UNIPROT:P33892; EMBL:U12467; NID:G117627; PIDN:AAA4635.1; PID:931  
A/Note: sequence extracted from NCBI backbone (NCBIN:132672, NCBI:P.132673)  
R/Kilma, R.; Cogilevina, M.; Bertani, I.; Zaccaria, P.; Brusch, C.V.  
submitted to the EMBL Data Library, September 1995  
A/Reference number: S62045  
A/Accession: S62050  
A/Molecule type: DNA  
A/Residues: 1-2672 <KLI>  
A/Cross-references: EMBL:X91837; NID:G117627; PID:e203621; PID:G1177633  
A/Experimental source: strain F11679

R.Brusch, C.V.; Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri, D.  
submitted to the Protein Sequence Database, May 1996

A:Reference number: S64183

A:Accession: S64212

A:Molecule type: DNA

A:Residues: 12672 <BRU>

A:Cross-references: EMBL:Z72717; NID:gl322820; PID:e43797; PID:gl322821; MIPS:YGL195W

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:GCN1

A:Cross-references: SGD:S0003163; MIPS:YGL195W

A:Map position: 7L

C:Keywords: translation regulation; transmembrane protein

F:68-85/Domain: transmembrane #status predicted <TM1>

F:216-232/Domain: transmembrane #status predicted <TM2>

F:929-945/Domain: transmembrane #status predicted <TM3>

F:993-1009/Domain: transmembrane #status predicted <TM4>

F:1704-1720/Domain: transmembrane #status predicted <TM5>

F:2021-2037/Domain: transmembrane #status predicted <TM6>

F:2272-2288/Domain: transmembrane #status predicted <TM7>

Query Match 29.6%; Score 64; DB 2; Length 2672;  
Best Local Similarity 34.2%; Pred. No. 38;

Matches 13; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

3 KEALIRLVALLLEBAEVINOKL--ASDPALRSKLVRLSS 40

778 KEQAKVKKLSKEQELVNEGLAKSAVRSHVSRIST 815

### RESULT 3

protoporphyrin IX magnesium chelatase (EC 4.99.1.-) chain chlI [validated] - *Synechocystis*

C:Species: *Synechocystis* sp.

C>Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 09-May-2004

C:Accession: T46868

R:Jensen, P.E.; Gibson, L.C.D.; Henningsen, K.W.; Hunter, C.N.

J. Biol. Chem. 271, 16662-16667, 1996

A:Title: Expression of the chlI, chlD, and chlH genes from the *Cyanobacterium synecocystis*

oltoporphyrin chelatase activity.

A:Reference number: Z24116; MUID:96279237; PMID:8663186

A:Accession: T46868

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-357 <JEN>

A:Cross-references: EMBL:U5144; NID:gl144564; PIDN:AAQ4138.1; PID:gl144565

A:Experimental source: strain PCC6803

C:Genetics:

A:Gene: chlI

C:Function:

A:Description: the insertion of Mg(2+) into protoporphyrin IX can only be catalyzed if a

C:Superfamily: magnesium chelatase, subunit ChlI

C:Keywords: lyase

Query Match 28.6%; Score 61.5; DB 2; Length 357;

Best Local Similarity 40.0%; Pred. No. 8.1; 20; Indels 3; Gaps 1;

Matches 18; Conservative 4; Mismatches 1;

1 TGKATIRLVALLLEBAEVINOKL--ASDPALRSKLVRLSS 42

51 TGKSTIRLADLLPEIEVANDPFPNSPSPDEPMSEEVIRVDS 95

### RESULT 4

protoporphyrin IX magnesium chelatase chain chlI-related protein - *Synechocystis* sp. (st

N:Alternate names: Mg-protoporphyrin IX chelatase; protein sir1030

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004

C:Accession: S75252

R:Kaneko, T.; Sato, S.; Kocani, H.; Tanaka, A.; Aamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimpou, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-116, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75252

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-357 <KAN>

A:Cross-references: UNIPROT:P51633; EMBL:D90904; GB:AB001339; NID:91652225; PIDN:BA117

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: chlI

C:Superfamily: magnesium chelatase, subunit ChlI

Query Match 28.6%; Score 61.5; DB 2; Length 357;

Best Local Similarity 40.0%; Pred. No. 8.1; 20; Indels 3; Gaps 1;

Matches 18; Conservative 4; Mismatches 1;

1 TGKATIRLVALLLEBAEVINOKL--ASDPALRSKLVRLSS 42

51 TGKSTIRLADLLPEIEVANDPFPNSPSPDEPMSEEVIRVDS 95

### RESULT 5

hypothetical protein alr7649 [imported] - *Nostoc* sp. (strain PCC 7120) plasmid pCC7120

C:Species: *Nostoc* sp. PCC 7120

A>Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004

C:Accession: AF2545

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuch

Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ar*

A:Reference number: AB1807; MUID:1535285; PMID:11759640

A:Accession: AF2545

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2581 <YUR>

A:Cross-references: UNIPROT:Q8Z563; GB:AF003602; PIDN:BA57292.1; PID:gl17134734; GSPDB

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr7649

A:Genome: plasmid

Query Match 28.4%; Score 61; DB 2; Length 2581;

Best Local Similarity 34.2%; Pred. No. 84;

Matches 13; Conservative 10; Mismatches 11; Indels 4; Gaps 1;

12 ALLEBAEVINOKL--ASDPALRSKLVRLSSDPFAH 45

310 AFLEQATLEBAEDPTVAEBLETRSRITATDSFIY 347

### RESULT 6

ABC transport system ATP-binding protein Cj0888c [imported] - *Campylobacter jejuni* (str

C:Species: *Campylobacter jejuni*

C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 16-Aug-2004

C:Accession: H81361

R:Parikh, J.; Wren, B.W.; Mungall, K.; Kexley, J.M.; Churcher, C.; Basham, D.; ChlI

C.W.; Quail, M.; Rajadream, M.A.; Rutherford, K.M.; VanYlief, A.; Whitehead, S.; Barr

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals h

A:Reference number: AB1250; MUID:20150912; PMID:10688204

A:Accession: H81361

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-643 <PAR>

A:Cross-references: UNIPROT:Q9PP43; GB:AL139076; GB:AL111168; NID:96960128; PIDN:CA573

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj0888c

C:Superfamily: ATP-binding cassette homology

Query Match 27.9%; Score 60; DB 2; Length 643;  
Best Local Similarity 35.0%; Pred. No. 24;  
Matches 14; Conservative 11; Mismatches 13; Indels 2; Gaps 1;

QY 3 KEAIRLVALLE--EAAVINOKLASPALSKVRLSS 40  
DB 84 KEAIRKQELVEIYNTLEVEYELHKKLEDPSSNDYLOKIDT 123

RESULT 7  
B82980  
probable transcription regulator PA5324 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: B82980  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lapidig, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; PMID:20437337; PMID:10984043  
A:Accession: B82980  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-356 <STO>  
A:Cross-references: UNIPROT:Q9HTN1; GB:AE004945; GB:AE004031; NID:G9951639; PIDN:AA60870  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA5324

Query Match 27.4%; Score 59; DB 2; Length 356;  
Best Local Similarity 33.3%; Pred. No. 16;  
Matches 15; Conservative 13; Mismatches 7; Indels 10; Gaps 1;

QY 7 LRRLLVALLBEAAVINOKL-----ASDPALSKVRLSSD 41  
DB 222 LRRQALLLEQESFOVHRKLDVQAVERRALARDLSGQIERISAE 266

RESULT 8  
AH1825  
protoporphyrin IX magnesium chelatase chain [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: AH1825  
R:Kanevo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Triguero  
Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; PMID:21595285; PMID:11759840  
A:Accession: AH1825  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-374 <XLR>  
A:Cross-references: UNIPROT:P58571; GB:BA000019; PIDN:BAH7676.1; PID:G17135130; GSPDB:Q  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: chl1  
C:Superfamily: magnesium chelatase, subunit Chl1

Query Match 27.2%; Score 58.5; DB 2; Length 374;  
Best Local Similarity 38.8%; Pred. No. 20;  
Matches 19; Conservative 4; Mismatches 23; Indels 3; Gaps 1;

QY 1 TGKEAIRLVALLEBEAAVINOKLASDPA---LRSKVRLSSDPAHL 46  
DB 51 TGKSTTIRALADLLPEIPVAVNDPFPNSDSDPLMSDEVKQSGTAIEI 99

RESULT 9

E75562  
Mutr/nudix family protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: E75562  
R:White, O.; Eisen, J.A.; Heldberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; PMID:20036896; PMID:10567266  
A:Accession: E75562  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-171 <WHI>  
A:Cross-references: UNIPROT:Q9RY71; GB:AE001871; GB:AE000513; NID:96457740; PIDN:AAF09  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0079  
A:Map position: 1

Query Match 27.0%; Score 58; DB 2; Length 171;  
Best Local Similarity 50.0%; Pred. No. 9.4;  
Matches 13; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 13 LLEEEAAVINOKLASDPALSKVRL 38  
DB 11 LVNERDEVGQILRTDPAIRKWRV 36

RESULT 10  
D69394  
phosphoribosylamine-glycine ligase (purf) homolog - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: D69394  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dods  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirsch, E.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes,  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archa  
A:Reference number: A69250; PMID:96049343; PMID:9389475  
A:Accession: D69394  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-470 <KLE>  
A:Cross-references: UNIPROT:Q29108; GB:AE001024; GB:AE000782; NID:G2689347; PIDN:AA890  
A:Superfamily: phosphoribosylamine-glycine ligase; phosphoribosylamine-glycine ligase  
F.2-467/Domain: phosphoribosylamine-glycine ligase homology <PGL>

Query Match 26.7%; Score 57.5; DB 2; Length 470;  
Best Local Similarity 34.9%; Pred. No. 33;  
Matches 15; Conservative 8; Mismatches 15; Indels 5; Gaps 1;

QY 4 EAIRLVALLEBEAAVINOKLA-----SDPALSKVRLSSD 41  
DB 302 KGVLYPVLIVAEENGELIPKYLEINVRDCPGAEAKLFRKSD 344

RESULT 11

B70612  
probable recD protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: B70612

R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genom

A:Reference number: A70500; MUID:98285987; PMID:9634230  
 A:Accession: B70612  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-575 <COL>  
 A:Cross-references: UNIPROT:P96919; GB:292772; GB:AL123456; NID:G3261722; PIDN:CA807118.  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: recD  
 C:Superfamily: exodeoxyribonuclease V 67K chain

Query Match 26.7%; Score 57.5; DB 2; Length 575;  
 Best Local Similarity 37.0%; Pred. No. 42;  
 Matches 17; Conservative 7; Mismatches 17; Indels 5; Gaps 1;

QY 1 TGKALIRLVALLIEEAIVINOKLASDPALRSKVLSSDSFAH 46  
 DB 174 TGKTTVARLALVAEOAE---LAGEPRPRLALAPGKAAAR 214

## RESULT 12

S52110  
 C:Species: Mus musculus (house mouse)  
 C:Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
 C:Accession: S52110; PH1578  
 R:Ota, A.; Ichinose, H.; Negatsu, T.  
 Biochim. Biophys. Acta 1260, 320-322, 1995  
 A:Title: Mouse sepiapterin reductase: an enzyme involved in the final step of tetrahydro  
 A:Reference number: S52110; MUID:95178553; PMID:7873607  
 A:Accession: S52110  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-261 <OTA>  
 A:Cross-references: UNIPROT:Q64105; GB:S77493; NID:G957229; PIDN:AB33611.1; PID:G957230  
 R:Meier, J.; Schott, K.; Werner, T.; Bacher, A.; Ziegler, I.  
 Exp. Cell Res. 204, 217-222, 1993  
 A:Title: Detection of a novel sepiapterin reductase mRNA: Assay of mRNA in various cells  
 A:Reference number: A49174; MUID:93178546; PMID:8440319  
 A:Accession: PH1578  
 A:Molecule type: mRNA  
 A:Residues: 209-255 <MAI>  
 C:Comment: This enzyme catalyzes the reduction of both C1' and C2' oxo group.  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 C:Keywords: oxidoreductase  
 F:9-200/Domains: short-chain alcohol dehydrogenase homology <SADH>

Query Match 26.5%; Score 57; DB 2; Length 261;  
 Best Local Similarity 39.3%; Pred. No. 20;  
 Matches 11; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 14 LEEBAIVINOKLASDPALRSKVLSSD 41  
 DB 202 LDNDMOOLARETSKDEPLRSKQKXSD 229

## RESULT 13

AC1740  
 C:Species: Listeria innocua  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AC1740  
 R:Glaser, P.; Frangoul, L.; Bachthieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioeckel  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussuguet, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 A:Title: Comparative genomics of *Listeria* species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AC1740  
 A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-504 <GLA>  
 A:Cross-references: UNIPROT:Q928R8; GB:AL592022; PIDN:CA697691.1; PID:G16414986; GSPDB  
 A:Experimental source: strain C11p1262  
 C:Genetics:  
 A:Gene: lin2464

Query Match 26.5%; Score 57; DB 2; Length 504;  
 Best Local Similarity 33.3%; Pred. No. 41;  
 Matches 17; Conservative 11; Mismatches 17; Indels 6; Gaps 1;

QY 1 TGKALIRLVALLIEEAIVINOKLASDPALRSKVLSSDSFAH 45  
 DB 147 TGKRYRVYIALILQAEYGVRIYPLDHEKMDIDFWATLNLKINIDAVAH 197

## RESULT 14

E72224  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 16-Aug-2004  
 C:Accession: E72224  
 R:Neison, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hick  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D  
 C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se  
 A:Reference number: A72200; MUID:9287316; PMID:10360571  
 A:Accession: E72224  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-259 <ARN>  
 A:Cross-references: UNIPROT:Q9X1Z1; GB:AE001808; GB:AE000512; NID:G4982233; PIDN:AA036  
 C:Genetics:  
 A:Experimental source: strain MSB8  
 C:Superfamily: ATP-binding cassette homology

Query Match 26.0%; Score 56; DB 2; Length 259;  
 Best Local Similarity 38.9%; Pred. No. 26;  
 Matches 14; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 TGKALIRLVALLIEEAIVINOKLASDPALRSKVL 36  
 DB 39 SGKTTLLKILAGLEAGELFLDGSPPADPPLRLKNV 74

## RESULT 15

AF2979  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 C:Accession: AF2979  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Mo  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCle  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: AF2979  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-412 <KOR>  
 A:Cross-references: UNIPROT:Q8UAD5; GB:AE008689; PIDN:AAU4452.1; PID:G17741836; GSPDB  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: acd  
 A:Map position: linear chromosome

Query Match 26.0%; Score 56; DB 2; Length 412;  
 Best Local Similarity 34.9%; Pred. No. 44;  
 Matches 15; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

us-10-092-750-4.rpr

Page 5

Search completed: November 10, 2004, 12:29:02  
Job time : 12.2994 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 10, 2004, 11:27:01 ; Search time 50.3302 Seconds  
(without alignments)  
525.871 Million cell updates/sec

Title: US-10-092-750-4

Perfect score: 215

Sequence: 1 TGGAATLRRLVALLEAEV.....SDPALRSKLVRLSDSPAH 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: uniprot\_spport.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	139	92.6	334 1 BCL2_HUMAN	Q9H09 homo sapien
2	173	80.5	255 2 Q9D3J3	Q9D3J3 mus musculu
3	106	49.3	187 2 Q8CHQ5	Q8CHQ5 mus musculu
4	67	31.2	105 2 Q7V7W8	Q7V7W8 prochloroco
5	67	31.2	830 2 Q6UXE9	Q6UXE9 homo sapien
6	67	31.2	830 2 AAO88750	AAO88750 homo sapi
7	66	30.7	833 2 Q9UP54	Q9UP54 homo sapien
8	66	30.7	1335 2 Q86VH1	Q86VH1 homo sapien
9	66	30.7	1401 2 Q86VH2	Q86VH2 homo sapien
10	65	30.2	1266 2 Q7PCK6	Q7PCK6 macaca faec
11	65	30.2	1394 2 Q7W6Z5	Q7W6Z5 rattus norv
12	64	29.8	2672 1 GCN1_YEAST	G31892 saccharomyc
13	63	29.3	468 2 Q81W72	Q81W72 homo sapien
14	62	28.8	504 2 Q71X64	Q71X64 listeria mo
15	62	28.8	504 2 AAT05101	AAT05101 listeria
16	61.5	28.6	357 1 CHL1_STYNY	PS1634 synecocyst
17	61	28.4	143 2 Q7W2Y9	Q7W2Y9 photorhabdu
18	61	28.4	722 2 Q6D7A6	Q6D7A6 erwinia car
19	61	28.4	2581 2 Q8Z853	Q8Z853 anabaena sp
20	60	27.9	278 2 Q88E83	Q88E83 pseudomonas
21	60	27.9	431 2 Q7N4J5	Q7N4J5 mycoplasma
22	60	27.9	643 2 Q9PP43	Q9PP43 campylobact
23	60	27.9	669 2 Q8BMB9	Q8BMB9 mus musculu
24	60	27.9	1394 2 Q7W6Z4	Q7W6Z4 mus musculu
25	59.5	27.7	302 2 Q74818	Q74818 geobacter s
26	59.5	27.7	302 2 AAR36406	AAR36406 geobacter
27	59.5	27.7	506 2 Q9NGY2	Q9NGY2 diffloraria
28	59	27.4	148 2 Q8G014	Q8G014 bruceella su
29	59	27.4	356 2 Q9HNT1	Q9HNT1 pseudomonas
30	59	27.4	429 2 Q71178	Q71178 brachydanio
31	59	27.4	565 2 Q7PON3	Q7PON3 chromobacte

32	59	27.4	749 2 Q6L433	Q6L433 solanum dem
33	59	27.4	749 2 AAT39950	AAT39950 solanum d
34	59	27.4	771 2 Q6L451	Q6L451 solanum dem
35	59	27.4	771 2 AAT39932	AAT39932 solanum d
36	58.5	27.2	223 2 Q9L218	Q9L218 streptomyces
37	58.5	27.2	338 1 CHL1_ANAVA	Q44498 anabaena va
38	58.5	27.2	374 1 CHL1_ANASP	PS6871 anabaena sp
39	58.5	27.2	790 2 Q92KX4	Q92KX4 rhizobium m
40	58.5	27.2	1479 2 Q9ATY5	Q9ATY5 arabidopsis
41	58.5	27.2	1522 2 Q9LRT2	Q9LRT2 arabidopsis
42	58	27.0	171 1 Y079_DEIRA	Q9Y71 deinococcus
43	58	27.0	430 2 Q89R53	Q89R53 bradyrhizob
44	58	27.0	744 2 Q6L466	Q6L466 solanum dem
45	58	27.0	744 2 AAT40519	AAT40519 solanum d

#### ALIGNMENTS

RESULT 1  
BCL2\_HUMAN  
ID BCL2\_HUMAN STANDARD: PRT; 334 AA.  
AC Q9H09; Q96196; Q9H08;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Bcl-2 related proline-rich protein (Bcl-2-like 12 protein).  
GN Name=BCL2L12; Synonyms=BPR;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=2295042; PubMed=1401436; DOI=10.1006/geno.2000.6455;  
RA Scottias A., Kyriakopoulos L., Yousef G.M., Ashworth L.K., Kwanle A.,  
RA Diamandis E.P.;  
RT "Molecular cloning, physical mapping, and expression analysis of a  
RT novel gene, BCL2L12, encoding a proline-rich protein with a highly  
RT conserved BH2 domain of the Bcl-2 family.";  
RL Genomics 72:217-221(2001).  
RN [2]  
RP SEQUENCE OF 62-334 FROM N.A. (ISOFORM 1).  
RC TISSUE=Uterus;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedlin T.B., Toshikiyuki S., Cavonius P., Pirange C.,  
RA Rana S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarnatone P.H.,  
RA Richards S., Kowley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey U., Helton E., Kettelman W., Madan A.C., Rodriques S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
RA Scherker A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SPICE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).  
RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-18;  
RA Hillman R.T., Green R.E., Brenner S.E.;  
RT "An unappreciated role for RNA surveillance.";  
RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).  
CC -!- ALTERNATIVE PRODUCTS;  
Event=Alternative splicing; Named isoforms=2;

```

CC      Name=1;
CC      Isoid=Q9HB09-1; Sequence=Displayed;
CC      Name=2;
CC      Isoid=Q9HB09-2; Sequence=VSP_000522, VSP_000523;
CC      Note=May be produced at very low levels due to a premature stop
CC      codon in the mRNA, leading to nonsense-mediated mRNA decay;
CC      -1- TISSUE SPECIFICITY: Expressed mainly in breast, thymus, prostate,
CC      fetal liver, colon, placenta, pancreas, small intestine, spinal
CC      cord, kidney, and bone marrow and to a lesser extent in many other
CC      tissues. Isoform 2 is primarily expressed in skeletal muscle.
CC      -1- SIMILARITY: Belongs to the Bcl-2 family.
CC      -1- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF289220; MAG29495.1; -
DR      EMBL; AF289220; MAG29496.1; -
DR      EMBL; BC007724; AAH07724.2; -
DR      Genew; HGNC:13787; BCL2L12.
DR      InterPro; IPR000712; Bcl2 BH.
DR      PROSITE; PS01258; BH2; FALSE NEG.
KW      Alternative splicing; Apoptosis.
FT      DOMAIN 311
FT      VARSPUBLIC 121 176
FT      SPQSEPTPTPLSRRLRCLPCSGRGRAAESRPRPSLPRLP
FT      CGLEPGPATPDPTFA -> PYSRLRCFGPAGAGPAA
FT      BISAQPRINGSPIIDREGSHTAAGPAGGGRSH (in
FT      isoform 2).
FT      /Frid=VSP_000522.
FT      /Frid=VSP_000523.
FT      /Frid=VSP_000523.
FT      P -> S (in Ref. 2).
FT      VARSPUBLIC 177 334
FT      MISSING (in isoform 2).
FT      /Frid=VSP_000523.
FT      /Frid=VSP_000523.
SQ      SEQUENCE 334 AA; 36821 MW; 5398E54C83E7CAB7 CRC64;

Query Match          92.6%; Score 199; DB 1; Length 334;
Best Local Similarity 95.7%; Pred. No. 4.7e-15;
Matches 44; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 TGEATLRRLVALLEEEAVYNOKLASDPALSKVLRSSDSFAHL 46
Db      204 TKEATLRRLVALLEEEAVYNOKLASDPALSKVLRSSDSFAHL 249

RESULT 2
Q9D3J3      PRELIMINARY;      PRT;      255 AA.
AC      Q9D3J3;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched
DE      library, clone:5430429M05 product:weakly similar to Bcl-2 RELATED
DE      PROLINE-RICH PROTEIN.
GN      Name=Bcl2l12;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Head;
RX      MEDLINE=99279253; PubMed=10349636;
RA      Carninci P., Hayashizaki Y.;
RT      "High-efficiency full-length cDNA cloning.";
RL      Meth. Enzymol. 303:19-44(1999).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Head;

```

```

RX      MEDLINE=21085660; PubMed=11217851;
RA      RIKEN FANTOM Consortium;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Head;
RA      The FANTOM Consortium.
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573(2002).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Head;
RX      MEDLINE=20499374; PubMed=11042159;
RA      Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA      Komori H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT      "Normalization and subtraction of cap-trapper-selected cDNAs to
RT      prepare full-length cDNA libraries for rapid discovery of new genes.";
RL      Genome Res. 10:1617-1630(2000).
RN      [5]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Head;
RX      MEDLINE=20330913; PubMed=11076861;
RA      Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA      Komori H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA      Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA      Yamamoto R., Matsunoto H., Sakauchi S., Ikegami T., Kashiwagi K.,
RA      Fujitake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA      Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunuma S., Kawai J.,
RA      Okazaki Y., Muramatsu M., Inoue Y., Kita T., Hayashizaki Y.;
RT      "RIKEN integrated sequence analysis (RISA) system-384-format
RT      sequencing pipeline with 384 multichipillary sequencer.";
RL      Genome Res. 10:1757-1771(2000).
RN      [6]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Head;
RA      Adachi U., Aizawa K., Akamita S., Akimura T., Arai A., Aono H.,
RA      Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA      Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirakawa T., Hori F.,
RA      Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA      Kawai J., Kojima Y., Komori H., Kouda M., Koya S., Kurihara C.,
RA      Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA      Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA      Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA      Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA      Teijima Y., Toya T., Yamamura T., Yaenunishi A., Yoshida K., Yoshino M.,
RA      Muramatsu M., Hayashizaki Y.;
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AK017362; BAB30709.1; -
DR      MGP; MGI:1922986; Bcl2l12.
SQ      SEQUENCE 255 AA; 27333 MW; 014F4C46294A51D4 CRC64;

Query Match          80.5%; Score 173; DB 2; Length 255;
Best Local Similarity 82.6%; Pred. No. 4e-12;
Matches 38; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 TGEATLRRLVALLEEEAVYNOKLASDPALSKVLRSSDSFAHL 46
Db      120 TKEATLRRLVALLEEEAVYNOKLASDPALSKVLRSSDSFAHL 165

RESULT 3
Q9CHQ5      PRELIMINARY;      PRT;      187 AA.
AC      Q9CHQ5;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Bcl2l12 protein.
GN      Name=Bcl2l12;
OS      Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheaffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shcherenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC039807; AAH39807.1; -.
SQ
SEQUENCE 187 AA; 20508 MW; 5F5C9CADFPD14702 CRC64;

Query Match 49.3%; Score 106; DB 2; Length 187;
Best Local Similarity 85.2%; Pred. No. 0.00022;
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

CY 1 TEKALRLVALLEBEAVINOKLAS 27
Db 120 TEKALRLVALLEBEAVINOKVTS 146

RESULT 4
07V7M8 PRELIMINARY; PRT; 105 AA.
AC 07V7M8;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical.
GN Ordered locus names=PMT0613;
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcales;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;
RA Roco G., Larimer F.W., Lamerdin U.E., Malfatti S., Chain P.,
RA Ahlgren N.A., Arrelano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.L., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinner E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL: EX572096; CAB20788.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 105 AA; 11989 MW; A91F0F4897B9EAF0 CRC64;

```

```

Query Match 31.2%; Score 67; DB 2; Length 105;
Best Local Similarity 39.4%; Pred. No. 4.5;
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

CY 3 KEALRLVALLEBEAVINOKLASDPALSKVRLSS 40
Db 426 KEALRLKALMOEKGESRLRSQALNEDIVRVS 463

RESULT 6
AAQ88750 PRELIMINARY; PRT; 830 AA.
AC AAQ88750;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DE 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE EOYK340.
GN UNQ340.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12975309;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hase P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vanden R., Watanabe C., Wiand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.L.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
DR EMBL: AY358384; AAQ88750.1; -.
SQ SEQUENCE 830 AA; 95028 MW; E6BE03248603B520 CRC64;

Query Match 39.2%; Score 67; DB 2; Length 830;
Best Local Similarity 39.5%; Pred. No. 40;
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

CY 3 KEALRLVALLEBEAVINOKLASDPALSKVRLSS 40
Db 426 KEALRLKALMOEKGESRLRSQALNEDIVRVS 463

RESULT 6
AAQ88750 PRELIMINARY; PRT; 830 AA.
AC AAQ88750;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DE 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE EOYK340.
GN UNQ340.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12975309;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hase P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vanden R., Watanabe C., Wiand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.L.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
DR EMBL: AY358384; AAQ88750.1; -.
SQ SEQUENCE 830 AA; 95028 MW; E6BE03248603B520 CRC64;

```



OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OX NCBI\_Taxid=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22730772; PubMed=12783626;  
 RA Oduru S., Campbell J.L., Karri S., Hendry W.J., Khan S.A.,  
 Williams S.C.;  
 RT "Gene discovery in the hamster: a comparative genomics approach for  
 BMC annotation by sequencing of hamster testis cDNAs.";  
 RL BMC Genomics 4:22-22(2003).  
 CC -1- MISCELLANEOUS: The sequence shown here is derived from an  
 EMBL/GenBank/DDJ third party annotation (TPA) entry.  
 CC EMBL; BK001055; DAA0313.1;  
 DR GO; GO:0005875; Cytoskeleton associated complex; IEA.  
 DR GO; GO:000524; F-ATP binding; IEA.  
 DR GO; GO:0003774; F-motor activity; IEA.  
 DR InterPro; IPR001752; Kinesin\_motor.  
 DR Pfam; PF00225; Kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; UNKNOWN\_1.  
 DR PROSITE; PS00657; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KM ATP-binding; Microtubule; Motor protein.  
 SQ SEQUENCE 1266 AA; 144456 MW; D558A8E6B5CCE6 CRC64;

## Query Match

Best Local Similarity 30.2%; Score 65; DB 2; Length 1266;  
 Matches 14; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

Qy 3 KEALIRLVALLIEEAIVINOKLASDPALSKVLRLSS 40  
 Db 927 REAIVSKKALLQEKSHLEKRLRSSQALNTDNLKIST 964

## RESULT 11

ID Q7M6Z5 PRELIMINARY; PRT; 1394 AA.  
 AC Q7M6Z5;  
 DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Kinesin-related protein KIF27A.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22730772; PubMed=12783626;  
 RA Oduru S., Campbell J.L., Karri S., Hendry W.J., Khan S.A.,  
 Williams S.C.;  
 RT "Gene discovery in the hamster: a comparative genomics approach for  
 BMC annotation by sequencing of hamster testis cDNAs.";  
 RL BMC Genomics 4:22-22(2003).  
 CC -1- MISCELLANEOUS: The sequence shown here is derived from an  
 EMBL/GenBank/DDJ third party annotation (TPA) entry.  
 CC EMBL; BK001055; DAA0313.1;  
 DR GO; GO:0005875; Cytoskeleton associated complex; IEA.  
 DR GO; GO:000524; F-ATP binding; IEA.  
 DR GO; GO:0003774; F-motor activity; IEA.  
 DR InterPro; IPR001752; Kinesin\_motor.  
 DR Pfam; PF00225; Kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; UNKNOWN\_1.  
 DR PROSITE; PS00657; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KM ATP-binding; Microtubule; Motor protein.  
 SQ SEQUENCE 1394 AA; 158879 MW; 78E6AA7FF01C7387 CRC64;

Query Match 30.2%; Score 65; DB 2; Length 1394;  
 Best Local Similarity 36.8%; Pred. No. 1.2e+02;  
 Matches 14; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

Qy 3 KEALIRLVALLIEEAIVINOKLASDPALSKVLRLSS 40  
 Db 952 REAIVSKKALLQEKSHLEKRLRSSQALNTDNLKIST 969

## RESULT 12

ID GCN1\_YEAST STANDARD; PRT; 2672 AA.  
 AC P33892;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Translational activator GCN1.  
 GN Name=GCN1; OrderedLocNames=YGL195W; ORFNames=G1318;  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_Taxid=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93268304; PubMed=8497269;  
 RA Matton M.J., Crouch D., Hinnebusch A.G.;  
 RT "GCN1, a translational activator of GCN4 in Saccharomyces cerevisiae,  
 is required for phosphorylation of eukaryotic translation initiation  
 factor 2 by protein kinase GCN2.";  
 RL Mol. Cell. Biol. 13:3541-3556(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=S288c; FTL679;  
 RC MEDLINE=97197971; PubMed=9046087;  
 RA Coglievina M., Klima R., Bertani I., Delneri D., Zaccaria P.,  
 Bruschi C.V.;  
 RT "Sequencing of a 40.5 kb fragment located on the left arm of  
 chromosome VII from Saccharomyces cerevisiae.";  
 RL Yeast 13:55-64(1997).  
 CC -1- FUNCTION: Translational activator of GCN4. May be involved in  
 sensing charged tRNA and stimulating the kinase activity of GCN2  
 in amino acid-starved cells. Required in vivo for the  
 phosphorylation of eIF-2-alpha on serine-52 by the protein kinase  
 GCN2.  
 CC -1- SUBUNIT: Component of a heteromeric complex that includes GCN1 and  
 GCN20.  
 CC -1- SIMILARITY: Contains 20 HEAT repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
 CC -----  
 DR EMBL; L12467; AAA34635.1; -  
 DR EMBL; X91837; CA62949.1; -  
 DR EMBL; Z72717; CA96907.1; -  
 DR PIR; A48126; A48126.  
 DR GenBank; U14243; -  
 DR SGD; S0003163; GCN1.  
 DR GO; GO:0005830; C:cytosolic ribosome (sensu Eukarya); IDA.  
 DR GO; GO:0006448; P:regulation of translational elongation; IMP.  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR000357; HEAT.  
 DR PROSITE; PS50077; HEAT\_REPEAT; 4.  
 KM Activator; Repeat; Translation regulation.  
 DR ACtive; Repeat; Translation regulation.  
 FT REPEAT 932 970 HEAT 1.  
 FT REPEAT 1030 1067 HEAT 2.  
 FT REPEAT 1099 1138 HEAT 3.  
 FT REPEAT 1243 1281 HEAT 4.  
 FT REPEAT 1284 1321 HEAT 5.  
 FT REPEAT 1405 1442 HEAT 6.  
 FT REPEAT 1484 1521 HEAT 7.  
 FT REPEAT 1523 1559 HEAT 8.

```

FT REPEAT 1561 1598 HEAT 9.
FT REPEAT 1603 1640 HEAT 10.
FT REPEAT 1641 1679 HEAT 11.
FT REPEAT 1721 1758 HEAT 12.
FT REPEAT 1760 1796 HEAT 13.
FT REPEAT 1862 1903 HEAT 14.
FT REPEAT 1905 1942 HEAT 15.
FT REPEAT 1947 1984 HEAT 16.
FT REPEAT 1985 2024 HEAT 17.
FT REPEAT 2097 2134 HEAT 18.
FT REPEAT 2290 2328 HEAT 19.
FT REPEAT 2347 2384 HEAT 20.
SQ SEQUENCE 2672 AA; 296693 MW; 980FDD03753E9D1C CRC64;

Query Match
Best Local Similarity 29.8%; Score 64; DB 1; Length 2672;
Matches 13; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 3 KEA1LRVALLLEEEAEVNOCLASDPALRSKLVLS 40
DB 778 KEQAKNVKLSKEQELVNEQLAKESAVRSHVSEIST 815

RESULT 13
ID 081W72 PRELIMINARY; PRT; 468 AA.
AC 081W72;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE LOC374654 protein (Fragment).
GN Name=LOC374654;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tomshyuk S., Carninci P., Prange C.,
RA Raha S.S., Loughlano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
RA Vallatton D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC040878; AAH40878.1; -
FT NON TER 1
SQ SEQUENCE 468 AA; 54032 MW; CE4172079C033B92 CRC64;

Query Match
Best Local Similarity 29.3%; Score 63; DB 2; Length 468;
Matches 15; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

```

```

QY 3 KEA1LRVALLLEEEAEVNOCLASDPALRSKLVLS 40
DB 64 REA1LAKKALMOEKTGTLEIKLRSSQALNEDIVRSS 101

RESULT 14
ID 071X64 PRELIMINARY; PRT; 504 AA.
AC 071X64;
DT 05-JUN-2004 (TrEMBLrel. 27, Created)
DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcriptional regulator, RofA family.
GN OrderedCusNames=LMOF2365_2335;
OS Listeria monocytogenes (serotype 4b / strain F2365).
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OC NCBI_TaxID=265669;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=15115801; DOI=10.1093/nar/gkh562;
RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
RA Peterson J.D., White O., Nelson W.C., Nieman W.C., Beaman M.J.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S.B., Khouri H.M., Fedorova N.,
RA Forberger H.A., Tran B., Kathariou S., Wonderling L.D., Uhlrich G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
borne pathogen Listeria monocytogenes reveal new insights into the
core genome components of this species.";
RL Nucleic Acids Res. 32:2386-2395(2004).
DR EMBL; AE017330; AAT05101.1; -
DR InterPro; IPR010735; RofA.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF07003; RofA; 1.
KM Complete proteome.
SQ SEQUENCE 504 AA; 59477 MW; B5468812813FE08C CRC64;

Query Match
Best Local Similarity 28.8%; Score 62; DB 2; Length 504;
Matches 17; Conservative 11; Mismatches 17; Indels 6; Gaps 1;

QY 1 TGEKAVRVYLLALQAEVGVLYPLDEEKQMTIDPMATLNLINIDILAH 197
DB 147 TGEKAVRVYLLALQAEVGVLYPLDEEKQMTIDPMATLNLINIDILAH 197

RESULT 15
ID AAT05101 PRELIMINARY; PRT; 504 AA.
AC AAT05101;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcriptional regulator, RofA family.
GN LMOF2365_2335.
OS Listeria monocytogenes str. 4b F2365.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria;
OC Listeria monocytogenes.
OC NCBI_TaxID=265669;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4b F2365;
RC PubMed=15115801;
RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
RA Peterson J., White O., Nelson W.C., Nieman W., Beaman M.J.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S., Khouri H., Fedorova N.,
RA Forberger H., Tran B., Kathariou S., Wonderling L.D., Uhlrich G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
borne pathogen Listeria monocytogenes reveal new insights into the

```

Fri Nov 12 14:55:35 2004

us-10-092-750-4.rup

Page 7

RT Core genome components of this species."  
RL Nucleic Acids Res. 32:2386-2395(2004).  
DR EMBL; AE017330; AAT05101.1; -.  
SQ SEQUENCE 504 AA; 59477 MW; B5468812813FE08C CRC64;  
Query Match 28.8%; Score 62; DB 2; Length 504;  
Best Local Similarity 33.3%; Pred. No. 91;  
Matches 17; Conservative 11; Mismatches 17; Indels 6; Gaps 1;  
QY 1 TGEATLRRLVALLLEEAEVI-----NOKLASDPALRSKIVRLSDSPAH 45  
DB 147 TGEKRVRYLIALQLQAEYGVLIYPLLDEEKQIMIDPFWATLNRINIDTLAH 197

Search completed: November 10, 2004, 12:26:45  
Job time : 53.3302 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 ; Search time 8.93518 Seconds  
(without alignments)  
222.664 Million cell updates/sec

Title: US-10-092-750-5

Perfect score: 153  
Sequence: 1 ORGMLYYQTEKYDLAIKDLKEALIQLRGN 30

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_Aa:\*  
1: /cgm2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgm2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgm2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
4: /cgm2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
5: /cgm2\_6/prodata/1/iaa/PCETUS\_COMB.pep:\*  
6: /cgm2\_6/prodata/1/iaa/backfillseq1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	147	96.1	526	4	US-09-538-092-962
2	56	36.6	387	4	US-09-410-132-2
3	54	35.3	34	3	US-09-301-978C-21
4	53	34.6	824	4	US-09-538-092-1041
5	51	33.3	371	4	US-09-538-092-548
6	51	33.3	332	4	US-09-134-000C-6542
7	49.5	32.4	122	4	US-09-513-999C-8056
8	49.5	32.4	221	3	US-08-297-431B-32
9	49.5	32.4	222	2	US-08-756-771-3
10	49.5	32.4	222	2	US-08-756-771-4
11	49.5	32.4	222	2	US-09-096-571-3
12	49.5	32.4	222	2	US-09-096-571-4
13	49.5	32.4	222	3	US-09-309-320-3
14	49.5	32.4	222	3	US-09-309-320-4
15	49	32.0	308	4	US-09-602-878A-62
16	48.5	31.7	122	4	US-09-270-767-31826
17	48.5	31.7	122	4	US-09-270-767-47043
18	48.5	31.7	220	3	US-08-297-431B-31
19	48.5	31.7	221	3	US-08-297-431B-32
20	48.5	31.7	221	3	US-08-297-431B-4
21	48.5	31.7	221	3	US-08-297-431B-6
22	48.5	31.7	221	3	US-08-297-431B-8
23	48.5	31.7	221	3	US-08-297-431B-10
24	48.5	31.7	221	3	US-08-297-431B-12
25	48.5	31.7	221	3	US-08-297-431B-14
26	48.5	31.7	221	3	US-08-297-431B-16
27	48.5	31.7	221	3	US-08-297-431B-18

28	48.5	31.7	221	3	US-08-297-431B-20	Sequence 20, Appl
29	48	31.4	222	2	US-08-756-771-5	Sequence 5, Appl
30	48	31.4	222	2	US-09-096-571-5	Sequence 5, Appl
31	48	31.4	222	3	US-09-309-320-5	Sequence 5, Appl
32	48	31.4	227	4	US-09-489-039A-13495	Sequence 13495, A
33	48	31.4	292	2	US-08-879-260-2	Sequence 2, Appl
34	48	31.4	292	4	US-09-581-001B-16	Sequence 16, Appl
35	48	31.4	909	4	US-09-248-786A-16165	Sequence 16165, A
36	47.5	31.0	120	4	US-09-442-099A-12	Sequence 12, Appl
37	47.5	31.0	120	4	US-09-612-342-12	Sequence 12, Appl
38	47.5	31.0	120	4	US-09-612-421A-12	Sequence 12, Appl
39	47.5	31.0	181	4	US-09-442-099A-10	Sequence 10, Appl
40	47.5	31.0	181	4	US-09-612-342-10	Sequence 10, Appl
41	47.5	31.0	181	4	US-09-612-421A-10	Sequence 10, Appl
42	47	30.7	63	3	US-08-894-626-1	Sequence 1, Appl
43	47	30.7	68	3	US-09-527-236A-22	Sequence 22, Appl
44	47	30.7	68	4	US-09-756-854-22	Sequence 22, Appl
45	47	30.7	70	4	US-09-159-277A-5	Sequence 5, Appl

#### ALIGNMENTS

```

RESULT 1
US-09-538-092-962
; Sequence 962, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glaxo, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538, 092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01/178, 965
; PRIOR APPLICATION NUMBER: 60/178, 965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 962
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P13878
US-09-538-092-962

Query Match          96.1%; Score 147; DB 4; Length 526;
Best Local Similarity 100.0%; Pred. No. 3.6e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ORGMLYYQTEKYDLAIKDLKEALIQLRGN 29
Db      76 ORGMLYYQTEKYDLAIKDLKEALIQLRGN 104

RESULT 2
US-09-410-132-2
; Sequence 132, Application US/09410132
; Patent No. 6509458
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel E.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Stephen C.
; TITLE OF INVENTION: NOVEL GENE EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 1703-021, US1
; CURRENT APPLICATION NUMBER: US/09/410,132
; CURRENT FILING DATE: 1999-09-30
; EARLIER APPLICATION NUMBER: 60/102,572
; EARLIER FILING DATE: 1998-09-30

```

EARLIER APPLICATION NUMBER: 60/146,584  
EARLIER FILING DATE: 1999-07-28  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 387  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-410-132-2

Query Match 36.6%; Score 56; DB 4; Length 387;  
Best Local Similarity 35.7%; Pred. No. 2.1;  
Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Cy 2 RGMYYQTEKYDIAIKDLKEALIQKRN 29  
Db 49 RGLIYVELGQYGFALDFKQALISRTN 76

RESULT 3  
US-09-301-978C-21  
Sequence 21, Application US/09301978C  
Patent No. 6392015  
GENERAL INFORMATION:  
APPLICANT: Pangniban, Antonio  
APPLICANT: Callahan, Mark A.  
TITLE OF INVENTION: Method of Identifying Modulators of HIV-1 VP1 and GAG  
FILE REFERENCE: 960296, 95335  
CURRENT APPLICATION NUMBER: US/09/301,978C  
CURRENT FILING DATE: 1999-04-29  
PRIOR APPLICATION NUMBER: 60/083,567  
PRIOR FILING DATE: 1998-04-30  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 21  
LENGTH: 34  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(34)  
OTHER INFORMATION: TPR3  
US-09-301-978C-21

Query Match 35.3%; Score 54; DB 3; Length 34;  
Best Local Similarity 50.0%; Pred. No. 0.21;  
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Cy 2 RGMYYQTEKYDIAIKDLKEAL 23  
Db 4 RAAAYCRLEQYDLAIDCRTL 25

RESULT 4  
US-09-538-092-1041  
Sequence 1041, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Glot, Loic  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuratSeqFormatter Version 0.9  
SEQ ID NO 1041  
LENGTH: 824

TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (0)..(0)  
OTHER INFORMATION: Polypeptide Accession Number P30260  
US-09-538-092-1041

Query Match 34.6%; Score 53; DB 4; Length 824;  
Best Local Similarity 47.6%; Pred. No. 15;  
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Cy 3 GMLYYQTEKYDIAIKDLKEAL 23  
Db 642 GMLYYQTEKYDIAIKDLKEAL 662

RESULT 5  
US-09-538-092-548  
Sequence 548, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Glot, Loic  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuratSeqFormatter Version 0.9  
SEQ ID NO 548  
LENGTH: 371  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (0)..(0)  
OTHER INFORMATION: Polypeptide Accession Number YLR216C  
US-09-538-092-548

Query Match 33.3%; Score 51; DB 4; Length 371;  
Best Local Similarity 40.9%; Pred. No. 12;  
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Cy 1 ORGMLYYQTEKYDIAIKDLKEA 22  
Db 313 RGLIYHVDITDMLNDLEMA 334

RESULT 6  
US-09-134-000C-6542  
Sequence 6542, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: Patentin Version 3.1  
SEQ ID NO 6542  
LENGTH: 392  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-6542

Query Match 33.3%; Score 51; DB 4; Length 392;  
Best Local Similarity 43.5%; Pred. No. 12;  
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 4 MLYQTEKIDLAIKDKALIQ 26  
DB 159 MKFKMKPKYQSLKDKESVNSL 181

RESULT 7  
US-09-513-999C-8056

Sequence 8056, Application US/09513999C  
Patent No. 6783961  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Duclert, A.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
Patent No. 6783961  
FILE REFERENCE: 59, US2, REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 8056  
LENGTH: 122  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-513-999C-8056

Query Match 32.4%; Score 49.5; DB 4; Length 122;  
Best Local Similarity 50.0%; Pred. No. 4.9;  
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 2 RGMLYQTEKIDLAIKDKALIQ 26  
DB 69 RALINLYASKNLYGKIDIKERALLDM 94

RESULT 8  
US-08-297-431B-32

Sequence 32, Application US/08297431B  
Patent No. 6136605  
GENERAL INFORMATION:  
APPLICANT: Paul, William E  
APPLICANT: Gulick, Andrew M  
APPLICANT: Manoharan, T Herbert  
APPLICANT: Puchalski, Ralph B  
APPLICANT: Kramer, Katharine  
APPLICANT: Masserman Wyeth W  
TITLE OF INVENTION: NOVEL GLUTATHIONE S-TRANSFERASE ISOFORMS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saul, Ewing, Remick & Saul, LLP  
STREET: 1500 Market Street, 38th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: US  
ZIP: 19102-2186  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/297,431B  
FILING DATE: August 26, 1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Reed, Janet E.

REGISTRATION NUMBER: 36252  
REFERENCE/DOCKET NUMBER: WARF F039  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-972-8386  
TELEFAX: 215-972-2292

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:  
LENGTH: 221 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-297-431B-32

Query Match 32.4%; Score 49.5; DB 3; Length 221;  
Best Local Similarity 50.0%; Pred. No. 10;  
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 2 RGMLYQTEKIDLAIKDKALIQ 26  
DB 68 RALINLYASKNLYGKIDIKERALLDM 93

RESULT 9  
US-08-756-771-3

Sequence 3, Application US/08756771  
Patent No. 5817497  
GENERAL INFORMATION:  
APPLICANT: Goli, Surya K.  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: A NOVEL GLUTATHIONE S-TRANSFERASE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/756,771  
FILING DATE: Herewith  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0162 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 825605  
US-08-756-771-3

Query Match 32.4%; Score 49.5; DB 2; Length 222;  
Best Local Similarity 50.0%; Pred. No. 10;  
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 2 RGMUYQTEKYDAIKDLKE-ALIQI 26  
DB 69 RALINVIASKYNLYGKDIERKALIDM 94

RESULT 10  
US-08-756-771-4  
Sequence 4, Application US/08756771  
Patent No. 5817497  
GENERAL INFORMATION:  
APPLICANT: Goll, Surya K.  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: A NOVEL GLUTATHIONE S-TRANSFERASE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/756,771  
FILING DATE: Herewith  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0162 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 259141  
US-08-756-771-4  
Query Match 32.4%; Score 49.5; DB 2; Length 222;  
Best Local Similarity 50.0%; Pred. No. 10;  
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;  
QY 2 RGMUYQTEKYDAIKDLKE-ALIQI 26  
DB 69 RALINVIASKYNLYGKDIERKALIDM 94  
RESULT 11  
US-09-096-571-3  
Sequence 3, Application US/09096571  
Patent No. 5976528  
GENERAL INFORMATION:  
APPLICANT: Goll, Surya K.  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: A NOVEL GLUTATHIONE S-TRANSFERASE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive

CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/096,571  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/756,771  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0162 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 825605  
US-09-096-571-3  
Query Match 32.4%; Score 49.5; DB 2; Length 222;  
Best Local Similarity 50.0%; Pred. No. 10;  
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;  
QY 2 RGMUYQTEKYDAIKDLKE-ALIQI 26  
DB 69 RALINVIASKYNLYGKDIERKALIDM 94  
RESULT 12  
US-09-096-571-4  
Sequence 4, Application US/09096571  
Patent No. 5976528  
GENERAL INFORMATION:  
APPLICANT: Goll, Surya K.  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: A NOVEL GLUTATHIONE S-TRANSFERASE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/096,571  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/756,771  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0162 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 259141  
US-09-096-571-4

Query Match 32.4%; Score 49.5; DB 2; Length 222;  
Best Local Similarity 50.0%; Pred. No. 10;  
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 2 RGMLYQTEKYDLAIKDLKE-ALIQL 26  
Db 69 RALNLYASKYNLYGDKIKERALLDM 94

RESULT 13  
US-09-309-320-3  
Sequence 3, Application US/09309320  
Patent No. 6248325  
GENERAL INFORMATION:  
APPLICANT: Goli, Surya K.  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: A NOVEL GLUTATHIONE S-TRANSFERASE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/309,320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/096,571  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0162 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 825605  
US-09-309-320-3

Query Match 32.4%; Score 49.5; DB 3; Length 222;  
Best Local Similarity 50.0%; Pred. No. 10;  
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 2 RGMLYQTEKYDLAIKDLKE-ALIQL 26  
Db 69 RALNLYASKYNLYGDKIKERALLDM 94

RESULT 14  
US-09-309-320-4  
Sequence 4, Application US/09309320  
Patent No. 6248325  
GENERAL INFORMATION:  
APPLICANT: Goli, Surya K.  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: A NOVEL GLUTATHIONE S-TRANSFERASE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/309,320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/096,571  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0162 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 259141  
US-09-309-320-4

Query Match 32.4%; Score 49.5; DB 3; Length 222;  
Best Local Similarity 50.0%; Pred. No. 10;  
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 2 RGMLYQTEKYDLAIKDLKE-ALIQL 26  
Db 69 RALNLYASKYNLYGDKIKERALLDM 94

RESULT 15  
US-09-602-787A-62  
Sequence 62, Application US/09602787A  
Patent No. 6696561  
GENERAL INFORMATION:  
APPLICANT: Pompeius, Mark  
APPLICANT: Krüger, Burkhard

APPLICANT: Schuder, Hartwig  
APPLICANT: Zelder, Oskar  
APPLICANT: Habernauer, Gregor  
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE  
TITLE OF INVENTION: TRANSPORT  
FILE REFERENCE: BGI-125CP  
CURRENT APPLICATION NUMBER: US/09/602,787A  
CURRENT FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: US98 60/141031  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: DE 19931454.3  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931478.0  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931563.9  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19932122.1  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932124.8  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932125.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932128.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932180.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932182.5  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932190.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932191.4  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932209.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932212.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932227.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932228.7  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932229.5  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932230.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932927.3  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933005.0  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933006.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19940764.9  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940765.7  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940766.5  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940830.0  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940831.9  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940832.7  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940833.5  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19941378.9  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941379.7  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941395.9  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19942077.7  
PRIOR FILING DATE: 1999-09-03

PRIOR APPLICATION NUMBER: DE 19942078.5  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942079.3  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942088.2  
PRIOR FILING DATE: 1999-09-03  
NUMBER OF SEQ ID NOS: 678  
SEQ ID NO 62  
LENGTH: 308  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-602-787A-62  
Query Match 32.0%; Score 49; DB 4; Length 308;  
Best Local Similarity 52.6%; Pred. No. 19;  
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 5 LYYQTEKYDAIKDLKEAL 23  
Db 57 LHLMEQODLPVXDLEAL 75

Search completed: November 10, 2004, 12:32:20  
Job time : 9.93518 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 ; Search time 28.9952 Seconds  
(without alignments)  
366.225 Million cell updates/sec

Title: US-10-092-750-5  
Sequence: 153  
1 QRGMLYYQTEKYDLAIKDKLKEALIQLRGNN 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35322586 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10D\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	153	100.0	30	US-10-092-750-5
2	147	96.1	526	US-10-418-036-20
3	62	40.5	987	US-10-283-122A-18445
4	58	37.9	152	US-10-425-114-61189
5	58	37.9	154	US-10-767-701-60819
6	58	37.9	155	US-10-425-114-42430
7	58	37.9	156	US-10-425-114-42430
8	58	37.9	157	US-10-425-114-70570
9	58	37.9	158	US-10-425-114-58967
10	58	37.9	159	US-10-425-114-68156
11	56	36.6	387	US-10-425-114-68156
12	55	35.9	179	US-10-437-963-114567
13	55	35.9	206	US-10-369-493-18850

14	54	35.3	34	9	US-09-301-978C-21	Sequence 21, Appl
15	54	35.3 <td>34</td> <td>13<td>US-10-090-378-21<td>Sequence 21, Appl</td></td></td>	34	13 <td>US-10-090-378-21<td>Sequence 21, Appl</td></td>	US-10-090-378-21 <td>Sequence 21, Appl</td>	Sequence 21, Appl
16	54	35.3 <td>180</td> <td>17<td>US-10-425-115-346616<td>Sequence 246616,</td></td></td>	180	17 <td>US-10-425-115-346616<td>Sequence 246616,</td></td>	US-10-425-115-346616 <td>Sequence 246616,</td>	Sequence 246616,
17	54	35.3 <td>446</td> <td>15<td>US-10-424-599-257381<td>Sequence 257381,</td></td></td>	446	15 <td>US-10-424-599-257381<td>Sequence 257381,</td></td>	US-10-424-599-257381 <td>Sequence 257381,</td>	Sequence 257381,
18	54	35.3 <td>491</td> <td>16<td>US-10-437-963-113247<td>Sequence 113247,</td></td></td>	491	16 <td>US-10-437-963-113247<td>Sequence 113247,</td></td>	US-10-437-963-113247 <td>Sequence 113247,</td>	Sequence 113247,
19	54	35.3 <td>607</td> <td>16<td>US-10-437-963-113245<td>Sequence 113245,</td></td></td>	607	16 <td>US-10-437-963-113245<td>Sequence 113245,</td></td>	US-10-437-963-113245 <td>Sequence 113245,</td>	Sequence 113245,
20	53	34.6 <td>851</td> <td>16<td>US-10-437-963-192117<td>Sequence 192117,</td></td></td>	851	16 <td>US-10-437-963-192117<td>Sequence 192117,</td></td>	US-10-437-963-192117 <td>Sequence 192117,</td>	Sequence 192117,
21	53	34.6 <td>1009</td> <td>17<td>US-10-425-115-346614<td>Sequence 346614,</td></td></td>	1009	17 <td>US-10-425-115-346614<td>Sequence 346614,</td></td>	US-10-425-115-346614 <td>Sequence 346614,</td>	Sequence 346614,
22	52	34.0 <td>169</td> <td>14<td>US-10-369-493-10893<td>Sequence 10893, A</td></td></td>	169	14 <td>US-10-369-493-10893<td>Sequence 10893, A</td></td>	US-10-369-493-10893 <td>Sequence 10893, A</td>	Sequence 10893, A
23	52	34.0 <td>245</td> <td>17<td>US-10-425-115-359760<td>Sequence 359760,</td></td></td>	245	17 <td>US-10-425-115-359760<td>Sequence 359760,</td></td>	US-10-425-115-359760 <td>Sequence 359760,</td>	Sequence 359760,
24	52	34.0 <td>1830</td> <td>16<td>US-10-437-963-189860<td>Sequence 189860,</td></td></td>	1830	16 <td>US-10-437-963-189860<td>Sequence 189860,</td></td>	US-10-437-963-189860 <td>Sequence 189860,</td>	Sequence 189860,
25	51.5	33.7 <td>659</td> <td>14<td>US-10-427-490-11<td>Sequence 11, Appl</td></td></td>	659	14 <td>US-10-427-490-11<td>Sequence 11, Appl</td></td>	US-10-427-490-11 <td>Sequence 11, Appl</td>	Sequence 11, Appl
26	51.5	33.7 <td>659</td> <td>14<td>US-10-356-456-11<td>Sequence 11, Appl</td></td></td>	659	14 <td>US-10-356-456-11<td>Sequence 11, Appl</td></td>	US-10-356-456-11 <td>Sequence 11, Appl</td>	Sequence 11, Appl
27	51.5	33.7 <td>788</td> <td>14<td>US-10-012-697-1523<td>Sequence 1523, Ap</td></td></td>	788	14 <td>US-10-012-697-1523<td>Sequence 1523, Ap</td></td>	US-10-012-697-1523 <td>Sequence 1523, Ap</td>	Sequence 1523, Ap
28	51.5	33.7 <td>839</td> <td>14<td>US-10-427-490-9<td>Sequence 9, Appl</td></td></td>	839	14 <td>US-10-427-490-9<td>Sequence 9, Appl</td></td>	US-10-427-490-9 <td>Sequence 9, Appl</td>	Sequence 9, Appl
29	51.5	33.7 <td>839</td> <td>14<td>US-10-327-490-10<td>Sequence 10, Appl</td></td></td>	839	14 <td>US-10-327-490-10<td>Sequence 10, Appl</td></td>	US-10-327-490-10 <td>Sequence 10, Appl</td>	Sequence 10, Appl
30	51.5	33.7 <td>839</td> <td>14<td>US-10-356-456-9<td>Sequence 9, Appl</td></td></td>	839	14 <td>US-10-356-456-9<td>Sequence 9, Appl</td></td>	US-10-356-456-9 <td>Sequence 9, Appl</td>	Sequence 9, Appl
31	51.5	33.7 <td>866</td> <td>14<td>US-10-356-456-10<td>Sequence 10, Appl</td></td></td>	866	14 <td>US-10-356-456-10<td>Sequence 10, Appl</td></td>	US-10-356-456-10 <td>Sequence 10, Appl</td>	Sequence 10, Appl
32	51.5	33.7 <td>866</td> <td>14<td>US-10-029-495-4<td>Sequence 4, Appl</td></td></td>	866	14 <td>US-10-029-495-4<td>Sequence 4, Appl</td></td>	US-10-029-495-4 <td>Sequence 4, Appl</td>	Sequence 4, Appl
33	51	33.3 <td>371</td> <td>16<td>US-10-451-467A-280<td>Sequence 280, App</td></td></td>	371	16 <td>US-10-451-467A-280<td>Sequence 280, App</td></td>	US-10-451-467A-280 <td>Sequence 280, App</td>	Sequence 280, App
34	51	33.3 <td>485</td> <td>14<td>US-10-369-493-12487<td>Sequence 12487, A</td></td></td>	485	14 <td>US-10-369-493-12487<td>Sequence 12487, A</td></td>	US-10-369-493-12487 <td>Sequence 12487, A</td>	Sequence 12487, A
35	51	33.3 <td>1733</td> <td>14<td>US-10-369-493-1564<td>Sequence 1564, Ap</td></td></td>	1733	14 <td>US-10-369-493-1564<td>Sequence 1564, Ap</td></td>	US-10-369-493-1564 <td>Sequence 1564, Ap</td>	Sequence 1564, Ap
36	50	32.7 <td>153</td> <td>17<td>US-10-425-115-203257<td>Sequence 203257,</td></td></td>	153	17 <td>US-10-425-115-203257<td>Sequence 203257,</td></td>	US-10-425-115-203257 <td>Sequence 203257,</td>	Sequence 203257,
37	50	32.7 <td>202</td> <td>17<td>US-10-425-115-276972<td>Sequence 276972,</td></td></td>	202	17 <td>US-10-425-115-276972<td>Sequence 276972,</td></td>	US-10-425-115-276972 <td>Sequence 276972,</td>	Sequence 276972,
38	50	32.7 <td>220</td> <td>17<td>US-10-298-638-126<td>Sequence 12, Appl</td></td></td>	220	17 <td>US-10-298-638-126<td>Sequence 12, Appl</td></td>	US-10-298-638-126 <td>Sequence 12, Appl</td>	Sequence 12, Appl
39	50	32.7 <td>222</td> <td>15<td>US-10-433-256-12<td>Sequence 12, Appl</td></td></td>	222	15 <td>US-10-433-256-12<td>Sequence 12, Appl</td></td>	US-10-433-256-12 <td>Sequence 12, Appl</td>	Sequence 12, Appl
40	50	32.7 <td>252</td> <td>17<td>US-10-425-115-281034<td>Sequence 15, Appl</td></td></td>	252	17 <td>US-10-425-115-281034<td>Sequence 15, Appl</td></td>	US-10-425-115-281034 <td>Sequence 15, Appl</td>	Sequence 15, Appl
41	50	32.7 <td>501</td> <td>15<td>US-10-427-224-15<td>Sequence 15, Appl</td></td></td>	501	15 <td>US-10-427-224-15<td>Sequence 15, Appl</td></td>	US-10-427-224-15 <td>Sequence 15, Appl</td>	Sequence 15, Appl
42	50	32.7 <td>579</td> <td>14<td>US-10-427-224-15<td>Sequence 15, Appl</td></td></td>	579	14 <td>US-10-427-224-15<td>Sequence 15, Appl</td></td>	US-10-427-224-15 <td>Sequence 15, Appl</td>	Sequence 15, Appl
43	50	32.7 <td>775</td> <td>15<td>US-10-369-493-11549<td>Sequence 11549, A</td></td></td>	775	15 <td>US-10-369-493-11549<td>Sequence 11549, A</td></td>	US-10-369-493-11549 <td>Sequence 11549, A</td>	Sequence 11549, A
44	50	32.7 <td>775</td> <td>15<td>US-10-425-114-53670<td>Sequence 53670, A</td></td></td>	775	15 <td>US-10-425-114-53670<td>Sequence 53670, A</td></td>	US-10-425-114-53670 <td>Sequence 53670, A</td>	Sequence 53670, A
45	50	32.7 <td>775</td> <td>17<td>US-10-425-115-323405<td>Sequence 323405,</td></td></td>	775	17 <td>US-10-425-115-323405<td>Sequence 323405,</td></td>	US-10-425-115-323405 <td>Sequence 323405,</td>	Sequence 323405,

#### ALIGNMENTS

RESULT 1  
US-10-092-750-5  
; Sequence 5, Application US/10092750  
; Publication No. US2003032517A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; APPLICANT: Alpin, Julia  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1  
; FILE REFERENCE: 50036/050002  
; CURRENT APPLICATION NUMBER: US/10/092, 750  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: US 60/274,526  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-750-5

Query Match 100.0%; Score 153; DB 14; Length 30;  
Best local Similarity 100.0%; Pred. No. 1.9e-14;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRGMLYYQTEKYDLAIKDKLKEALIQLRGNN 30  
DB 1 QRGMLYYQTEKYDLAIKDKLKEALIQLRGNN 30

RESULT 2  
US-10-418-036-20  
; Sequence 20, Application US/10418036  
; Publication No. US20030225117A1  
; GENERAL INFORMATION:

```

; APPLICANT: Gronberg, Alvar
; APPLICANT: Wikstrom, Per
; TITLE OF INVENTION: NEW USE
; FILE REFERENCE: 13425-11001
; CURRENT APPLICATION NUMBER: US/10/418,036
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: SE 0201152-6
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/410,626
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-418-036-20

Query Match          96.1%; Score 147; DB 14; Length 526;
Best Local Similarity 100.0%; Pred. No. 3,5e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ORGMLYYQTEKYDLAIKDKKALIQLRGN 29
Db      76  ORGMLYYQTEKYDLAIKDKKALIQLRGN 104

RESULT 3
US-10-282-122A-48445
; Sequence 48445, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangou
; APPLICANT: Zamudio, Carlos
; APPLICANT: Maldino, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: BLITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48445
; LENGTH: 987
```

```

; TYPE: PRT
; ORGANISM: Bacteroides fragilis
US-10-282-122A-48445

Query Match          40.5%; Score 62; DB 15; Length 987;
Best Local Similarity 42.9%; Pred. No. 7.5;
Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY      3  GMLYYQTEKYDLAIKDKKALIQLRGN 30
Db      648  GLIYYQTEKYDLAIKDKKALIQLRGN 675

RESULT 4
US-10-425-114-61189
; Sequence 61189, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61189
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3593-003-D12_FLI.pep
US-10-425-114-61189

Query Match          37.9%; Score 58; DB 15; Length 152;
Best Local Similarity 52.2%; Pred. No. 3.3;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      1  ORGMLYYQTEKYDLAIKDKKAL 23
Db      109  ORGMLYYQTEKYDLAIKDKKAL 131

RESULT 5
US-10-767-701-60819
; Sequence 60819, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 60819
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 9297924.pep
US-10-767-701-60819

Query Match          37.9%; Score 58; DB 16; Length 154;
Best Local Similarity 52.2%; Pred. No. 3.3;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
```



QY 1 ORGMLYYQTEKYDLAIKDLKEAL 23  
DB 111 QRALYESTEKYRLGAEDRLVL 133

## RESULT 6

US-10-425-114-42430  
; Sequence 42430, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yinhua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 42430  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3136-032-F4\_FLI.pep  
US-10-425-114-42430

Query Match 37.9%; Score 58; DB 15; Length 260;  
Best Local Similarity 52.2%; Pred. No. 6;

Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 ORGMLYYQTEKYDLAIKDLKEAL 23

DB 217 QRALYESTEKYRLGAEDRLVL 239

## RESULT 7

US-10-425-114-70570  
; Sequence 70570, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yinhua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 70570  
; LENGTH: 403  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700172224\_FLI.pep  
US-10-425-114-70570

Query Match 37.9%; Score 58; DB 15; Length 403;  
Best Local Similarity 52.2%; Pred. No. 9;

Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 ORGMLYYQTEKYDLAIKDLKEAL 23

DB 360 QRALYESTEKYRLGAEDRLVL 382

RESULT 8  
US-10-425-115-354212  
; Sequence 354212, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yinhua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 354212  
; LENGTH: 494  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(494)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MFT4577\_86212C.1.pep  
US-10-425-115-354212

Query Match 37.9%; Score 58; DB 17; Length 494;  
Best Local Similarity 52.2%; Pred. No. 13;

Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 ORGMLYYQTEKYDLAIKDLKEAL 23

DB 451 QRALYESTEKYRLGAEDRLVL 473

## RESULT 9

US-10-425-114-58967  
; Sequence 58967, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yinhua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 58967  
; LENGTH: 519  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700332031\_FLI.pep  
US-10-425-114-58967

Query Match 37.9%; Score 58; DB 15; Length 519;  
Best Local Similarity 52.2%; Pred. No. 13;

Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 ORGMLYYQTEKYDLAIKDLKEAL 23

DB 476 QRALYESTEKYRLGAEDRLVL 498

## RESULT 10

US-10-425-114-68156  
; Sequence 68156, Application US/10425114

```
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 68156
LENGTH: 525
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700051726_FLI.pep
US-10-425-114-68156

Query Match          37.9%; Score 58; DB 15; Length 525;
Best Local Similarity 52.2%; Pred. No. 13;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 RGMLYYQTEKYDLAIKDLKEAL 23
Db 482 QRALTYESTEKYRLGAEDELRLVL 504

RESULT 11
US-10-334-561A-2
Sequence 2, Application US/10334561A
Publication No. US20030211520A1
GENERAL INFORMATION:
APPLICANT: Afari, Daniel E.
APPLICANT: Hubert, Rene S.
APPLICANT: Mitchell, Stephen C.
TITLE OF INVENTION: NOVEL GENE EXPRESSED IN PROSTATE CANCER
FILE REFERENCE: 511582003110
CURRENT APPLICATION NUMBER: US/10/334,561A
CURRENT FILING DATE: 2002-12-30
PRIOR APPLICATION NUMBER: 09/410,132
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 60/146,584
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/102,572
PRIOR FILING DATE: 1998-09-30
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 387
TYPE: PRT
ORGANISM: Homo sapiens
US-10-334-561A-2

Query Match          36.6%; Score 56; DB 14; Length 387;
Best Local Similarity 35.7%; Pred. No. 18;
Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 2 RGMLYYQTEKYDLAIKDLKEAL 29
Db 49 RGLIYVELGQYGFALDEPKALISRTN 76

RESULT 12
US-10-437-963-114567
Sequence 21, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
```

```
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 114567
LENGTH: 179
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_18245C.1.pep
US-10-437-963-114567

Query Match          35.9%; Score 55; DB 16; Length 179;
Best Local Similarity 54.5%; Pred. No. 10;
Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 RGMLYYQTEKYDLAIKDLKEAL 23
Db 30 RGSYDRLGKRYDLAIQDYTKAL 51

RESULT 13
US-10-369-493-18850
Sequence 21, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 18850
LENGTH: 206
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-369-493-18850

Query Match          35.9%; Score 55; DB 14; Length 206;
Best Local Similarity 43.5%; Pred. No. 12;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 1 RGMLYYQTEKYDLAIKDLKEAL 23
Db 16 KRGLSYQDLQDYEAALADYSQAI 38

RESULT 14
US-09-301-978C-21
Sequence 21, Application US/09301978C
Publication No. US2002022221A1
GENERAL INFORMATION:
APPLICANT: Pangamban, Antonito
APPLICANT: Callahan, Mark A.
TITLE OF INVENTION: Method of Identifying Modulators of HIV-1 VPU and GAG
TITLE OF INVENTION: Interaction with U Binding Protein (UBP)
FILE REFERENCE: 960296.95335
CURRENT APPLICATION NUMBER: US/09/301,978C
```

```

:
: CURRENT FILING DATE: 1999-04-29
: PRIOR APPLICATION NUMBER: 60/083,566
: PRIOR FILING DATE: 1998-04-30
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 21
: LENGTH: 34
: TYPE: prt
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: PEPTIDE
: LOCATION: (1)..(34)
: OTHER INFORMATION: TRP3
US-09-301-978C-21

```

Query Match	35.3%	Score 54	DB 9	Length 34
Best Local Similarity	50.0%	Pred. No. 2.2		
Matches 11, Conservative	4	Mismatches 7	Indels 0	Gaps 0

QY	2	RGMLYYQTEKXNDLAIKDLKEAL	23
		: : : : : : : : : : :	
Db	4	RAAAYCRLEQYDLAIQDRTAL	25

```

RESULT 15
US-10-090-378-21
; Sequence 21, Application US/10090378
; Publication No. US20020115830a1
; GENERAL INFORMATION:
; APPLICANT: Pangamban, Antonito
; APPLICANT: Callahan, Mark A.
; TITLE OF INVENTION: Method of Identifying Modulators of HIV-1 VPU and GAG
; FILE REFERENCE: 960296.9535
; CURRENT APPLICATION NUMBER: US/10/090.378
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US/09/301.978C
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083,567
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(34)
; OTHER INFORMATION: TPR3
US-10-090-378-21

```

Query Match	35.3%	Score 54;	DB 13;	Length 34;
Best Local Similarity	50.0%;	Pred. No. 2.2;		
Matches 11;	Conservative 4;	Mismatches 7;	Indels 0;	Gaps 0;

DY      2 RGM<sup>L</sup><sup>T</sup>YYQTEKYDLAIKDLEAL 23  
         | : : : ||| : : |  
Db    4 RAAAYCRLEGYDIAIQDCRTAL 25

Search completed: November 10, 2004, 16:35:41  
Job time : 30.9352 secs

This Page Blank (uspto)

This Page Blank (uspto)

This Page Blank (uspto)

Fri Nov 12 14:55:46 2004

us-10-092-750-5.rpr

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:36:51 ; Search time 6.06481 Seconds  
(without alignments)  
475.942 Million cell updates/sec

Title: US-10-092-750-5

Perfect score: 153  
Sequence: 1 QRGMLYYQTEKYDLAIKDLKEALIQLRGNN 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR79:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	96.1	526	2	A34855 67K neutrophil oxi
2	62	40.5	174	2	AG2234 hypothetical prote
3	58.5	38.2	320	2	S76422 hypothetical prote
4	55	35.9	362	2	T31715 hypothetical prote
5	55	35.9	547	2	AE1884 hypothetical prote
6	54	35.3	337	2	TI6889 hypothetical prote
7	54	35.3	363	2	S76156 hypothetical prote
8	53.5	35.0	222	1	XUR7G glutathione transf
9	53.5	35.0	222	2	A25653 glutathione transf
10	53.5	35.0	223	1	A27848 glutathione transf
11	53	34.6	342	2	E70463 conserved hypotnet
12	53	34.6	824	2	I52835 H-NUC - human
13	52	34.0	217	2	B70330 hypothethical prote
14	52	34.0	317	2	TI4686 hypothethical prote
15	52	34.0	361	2	TI4923 hypothethical prote
16	52	34.0	774	2	T03919 hypothethical prote
17	51.5	33.7	151	2	S37006 transposase (clone
18	51.5	33.7	708	2	A47176 probable transmem
19	51	33.3	371	2	A48567 cyclophilin-like p
20	51	33.3	406	2	S36567 transcriptions acti
21	51	33.3	1733	1	RNBV2L DNA-directed RNA p
22	50.5	33.0	490	1	C57150 NADP-reducing hydr
23	50	32.7	272	2	AH1809 hypothethical prote
24	50	32.7	1115	2	B84476 probable TPR repea
25	49.5	32.4	191	2	S29657 glutathione trans
26	49.5	32.4	191	2	S29658 glutathione trans
27	49.5	32.4	221	2	S43431 glutathione trans
28	49.5	32.4	222	2	A49365 glutathione trans
29	49.5	32.4	222	2	S24330 glutathione trans

30	49.5	32.4	222	2	A56666 glutathione trans
31	49.5	32.4	229	2	S19734 glutathione trans
32	49.5	32.4	980	2	AD1922 methyl-accepting c
33	49	32.0	127	2	C81292 hypothethical prote
34	49	32.0	318	2	B36972 aad 5'-region hypo
35	49	32.0	318	2	S33433 hypothethical prote
36	49	32.0	1195	2	E96615 hypothethical prote
37	49	32.0	2166	2	G70163 hypothethical prote
38	48.5	31.7	221	1	A26753 glutathione trans
39	48.5	31.7	454	2	C90647 poly(A) polymerase
40	48.5	31.7	454	2	G64737 polynucleotide ade
41	48.5	31.7	454	2	C85498 poly(A) polymerase
42	48	31.4	183	2	A99262 hypothethical prote
43	48	31.4	219	2	A70314 deoxyribose-phosph
44	48	31.4	222	1	XUR78C glutathione trans
45	48	31.4	222	2	S27234 glutathione trans

## ALIGNMENTS

### RESULT 1

A34855  
67K neutrophil oxidase factor - human  
C/Species: Homo sapiens (man)  
C/Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 09-Jul-2004  
C/Accession: A34855  
R/Reto, T.L., Lomax, K.J., Volpp, B.D., Nunoi, H., Sechler, J.M.G., Nauseef, W.M., Clat  
Science 248, 727-729, 1990  
A/Title: Cloning of a 67-kD neutrophil oxidase factor with similarity to a noncatalytic  
A/Reference number: A34855; MUID:90239568; PMID:1692159  
A/Accession: A34855  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-526 <LST>  
A/Cross-references: UNIPROT:P19878; GB:M32011; NID:g189267; PIDN:AAA36379.1; PID:g18926  
C/Genetics:  
A/Gene: GDB:NCF2  
A/Cross-references: GDB:120223; OMIM:233710  
A/Map position: 1q25-1q25  
F/37-70/Domain: tetratricopeptide repeat homology <TT1>  
F/71-104/Domain: tetratricopeptide repeat homology <TT2>  
F/247-294/Domain: SH3 homology <SH31>  
F/464-511/Domain: SH3 homology <SH32>

Query Match 96.1% Score 147; DB 2; Length 526;  
Best local similarity 100.0% Pred. No. 2, 2e-12;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 QRGMLYYQTEKYDLAIKDLKEALIQLRGN 29  
DB 76 QRGMLYYQTEKYDLAIKDLKEALIQLRGN 104

### RESULT 2

AG2234  
Hypothetical protein alr3430 [imported] - Nostoc sp. (strain PCC 7120)  
C/Species: Nostoc sp. PCC 7120  
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Arabaena sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C/Accession: AG2234  
R/Kaneko, T., Nakamura, Y., Molk, C.P., Kuritz, T., Sasamoto, S., Watanabe, A., Iriqui  
Nakazaki, N., Shimo, S., Sugimoto, M., Takazawa, M., Yamada, M., Yasuda, M., Tabata,  
DNA Res. 8, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A  
A/Reference number: AB1807; MUID:21595285; PMID:11759840  
A/Accession: AG2234  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-174 <KUR>  
A/Cross-references: UNIPROT:Q8YRL6; GB:BA000019; PIDN:BA075129.1; PID:g17132563; GSPDB  
A/Experimental source: strain PCC 7120  
C/Genetics:

A/Gene: alr3430  
C/Superfamily: unassigned tetratricopeptide repeat proteins; tetratricopeptide repeat hc

Query Match  
Best Local Similarity 40.5%; Score 62; DB 2; Length 174;  
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 GMLYYQTEKYDLAIKDLKEAL 23  
DB 98 GYAFQSQYDIAIKDYKEAL 118

RESULT 3  
S76422  
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)  
C/Species: *Synechocystis* sp.  
A/Variety: PCC 6803  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: S76422  
R/Kaneke, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996  
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.  
A/Reference number: S74322; MUID:97061201; PMID:8905231  
A/Accession: S76422  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-320 <KAN>  
A/Cross-references: UNIPROT:P74450; EMBL:D90915; GB:AB001339; NID:g1653604; PIDN:BA1855  
A/Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match  
Best Local Similarity 38.2%; Score 58.5; DB 2; Length 320;  
Matches 13; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

QY 1 QRGMLYYQTEKYDLAIKDLKEAL 29  
DB 190 QRGSLHYRLDVAEAVRDCTEA-IRLRD 217

RESULT 4  
T31715  
hypothetical protein F44E7.5 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 04-Mar-2000  
C/Accession: T31715  
R/Du, Z.; Scheet, P.; Delehaunty, A.  
submitted to the EMBL Data Library, July 1997  
A/Description: The sequence of *C. elegans* cosmid F44E7.  
A/Reference number: Z21073  
A/Accession: T31715  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-362 <DUZ>  
A/Cross-references: EMBL:AF016421; PIDN:AAC25790.1; GSPDB:GN00023; CESP:F44E7.5  
A/Experimental source: strain Bristol N2; clone F44E7  
A/Genetics:  
A/Gene: CESP:F44E7.5  
A/Map position: 5  
A/Introns: 44/3; 83/2; 195/3  
C/Superfamily: *Caenorhabditis elegans* hypothetical protein F44E7.5

Query Match  
Best Local Similarity 35.9%; Score 55; DB 2; Length 362;  
Matches 16; Conservative 5; Mismatches 9; Indels 10; Gaps 2;

QY 1 QRGMLYYQ-----TEKYDLAIKDLKE-----ALIQLRGNN 30  
DB 80 RRGRAHFHSIFCLTDKIDLAIVNLLLEFAKELILFRNN 119

RESULT 5

AE1884  
hypothetical protein alr0622 [imported] - *Nostoc* sp. (strain PCC 7120)  
C/Species: *Nostoc* sp. PCC 7120  
A/Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C/Accession: AE1884  
R/Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchihara, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120  
A/Reference number: AB1807; MUID:21595285; PMID:11759840  
A/Accession: AE1884  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-547 <KUR>  
A/Cross-references: UNIPROT:Q8YZ65; GB:BA000013; PIDN:BAE72580.1; PID:g17129968; GSPDB  
A/Experimental source: strain PCC 7120  
A/Genetics:  
A/Gene: alr0622

Query Match  
Best Local Similarity 35.9%; Score 55; DB 2; Length 547;  
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 QRGMLYYQTEKYDLAIKDLKEAL 23  
DB 292 KRGLSYQLDVAEAAIDYQAI 314

RESULT 6  
T16689  
hypothetical protein R05F9.10 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 16-Aug-2004  
C/Accession: T16689  
R/Hallsworth, K.  
submitted to the EMBL Data Library, November 1995  
A/Description: The sequence of *C. elegans* cosmid R05F9.  
A/Reference number: Z18559  
A/Accession: T16689  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-337 <HAL>  
A/Cross-references: UNIPROT:Q21746; EMBL:U41533; NID:g1109807; PID:g1109816; PIDN:AAA83  
A/Genetics:  
A/Gene: CESP:R05F9.10  
A/Introns: 53/1; 282/3  
C/Superfamily: tetratricopeptide repeat homology

Query Match  
Best Local Similarity 35.3%; Score 54; DB 2; Length 337;  
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 RGMYYQTEKYDLAIKDLKEAL 23  
DB 144 RAAAYCRSLQYDLAIQDCRTAL 165

RESULT 7  
S76156  
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)  
C/Species: *Synechocystis* sp.  
A/Variety: PCC 6803  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 16-Aug-2004  
C/Accession: S76156  
R/Kaneke, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996  
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.  
A/Reference number: S74322; MUID:97061201; PMID:8905231  
A/Accession: S76156  
A/Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-363 <KAN>  
A:Cross-references: UNIPROT:P74321; EMBL:D09014; GB:AB001339; NID:G1653477; PIRN:BA11841  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: tetratricopeptide repeat homology  
F:159-192/Domain: tetratricopeptide repeat homology <TT1>  
F:193-226/Domain: tetratricopeptide repeat homology <TT2>  
F:230-263/Domain: tetratricopeptide repeat homology <TT3>  
F:264-297/Domain: tetratricopeptide repeat homology <TT4>  
F:298-331/Domain: tetratricopeptide repeat homology <TT5>

Query Match 35.3%; Score 54; DB 2; Length 363;  
Best Local Similarity 47.8%; Pred. No. 10;  
Matches 11; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 RGMUYQTEKYDLAIQDLKEAL 24  
DB 270 RGLAYVMQGIKXKXIDFSDALI 292

## RESULT 8

## XJRTG

glutathione transferase (EC 2.5.1.18) class alpha chain Ya1 - rat

N:Alternate names: glutathione S-alkyltransferase; glutathione S-aryltransferase; glutathione S-transferase

C:Species: Rattus norvegicus (Norway rat)

C>Date: 18-Aug-1982 #sequence revision 28-May-1986 #text\_change 07-May-1999

C:Accession: A92479; A92370; S64679; A00591

R:Ref: H.C.U.; Li, N.; Weiss, M.J.; Reddy, C.C.; Tu, C.P.D.

J: Biol. Chem. 259, 5536-5542, 1984

A:Title: The nucleotide sequence of a rat liver glutathione S-transferase subunit cDNA

A:Reference number: A92479; MUID:84185691; PMID:6201485

A:Accession: A92479

A:Molecule type: mRNA

A:Residues: 1-222 <LAI>

A:Experimental source: clone pGTR261, liver

R:Kajiyak, J.E.; Taylor, J.M.

J: Biol. Chem. 257, 523-530, 1982

A:Title: Rat glutathione S-transferase. Cloning of double-stranded cDNA and induction of

A:Reference number: A92370; MUID:82075944; PMID:6273441

A:Accession: A92370

A:Molecule type: mRNA

A:Residues: 46-197 <KAL>

R:Reh, H.I.; Lee, J.Y.; Tsai, S.P.; Hsieh, C.H.; Tam, M.F.

Biochem. J. 314, 1017-1025, 1996

A:Title: Rat kidney glutathione S-transferase 1 subunits have C-terminal truncations.

A:Reference number: S64679; MUID:96177880; PMID:8615753

A:Accession: S64679

A:Molecule type: protein

A:Residues: 32-207 'V', 209-222 <YEH>

A:Experimental source: kidney, liver

C:Comment: Glutathione transferases are multifunctional dimers of identical or similar

m-independent glutathione peroxidase activity toward organic hydroperoxides.

C:Comment: In addition to its enzymatic activity, the homodimer of Ya chains, called 119

my mammalian tissues.

C:Superfamily: glutathione transferase

C:Keywords: acetylated amino end; dimer; kidney; liver; transferase

F:2-222/Product: glutathione transferase 1, 25.5K (liver, kidney) #status experimental

F:2-221/Product: glutathione transferase 1, 25.3K (liver, kidney) #status experimental

F:2-218/Product: glutathione transferase 1, 24.9K (kidney) #status experimental <GTS>

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted

Query Match 35.0%; Score 53.5; DB 1; Length 222;  
Best Local Similarity 53.8%; Pred. No. 6.9;

Matches 14; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 2 RGMUYQTEKYDLAIQDLKE-ALIQ 26  
DB 69 RALNYIATKRYDYGDMKERALIDM 94

## RESULT 9

## A26653

glutathione transferase (EC 2.5.1.18) class alpha chain Ya2, hepatic - rat

N:Alternate names: GSH transferase chain Ya; glutathione transferase 1'  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 01-Feb-1988 #sequence revision 01-Feb-1988 #text\_change 09-Jul-2004  
A:Accession: A26653; I51831; S03358; S78477; S10556; I52350; A93438  
R:Telakowski-Hopkins, C.A.; Rothkopf, G.S.; Pickett, C.B.  
Proc. Natl. Acad. Sci. U.S.A. 83, 9393-9397, 1986  
A:Title: Structural analysis of a rat liver glutathione S-transferase Ya gene.

A:Reference number: A26653; MUID:87092258; PMID:3025841

A:Accession: A26653

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-222 <TEU>

A:Cross-references: UNIPROT:P04903; GB:M14991; NID:G204524; PIDN:AAA41295.1; PID:G2045

R:Pickett, C.B.; Telakowski-Hopkins, C.A.; Ding, G.J.

Adv. Exp. Med. Biol. 197, 185-193, 1986

A:Title: Expression and sequence analysis of rat liver glutathione S-transferase genes

A:Reference number: I51831; MUID:87022415; PMID:3766257

A:Accession: I51831

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-222 <RES>

A:Cross-references: GB:M25891; NID:G204507; PIDN:AAA41290.1; PID:G204508

R:Wang, R.W.; Pickett, C.B.; Lu, A.Y.H.

Arch. Biochem. Biophys. 269, 536-543, 1989

A:Title: Expression of a cDNA encoding a rat liver glutathione S-transferase Ya subunit

A:Reference number: S03358; MUID:89149100; PMID:2645828

A:Accession: S03358

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-51 <MAN>

A:Cross-references: GB:M27446; NID:G294563; PIDN:AAA41291.1; PID:G294564

A:Accession: S78477

A:Molecule type: protein

A:Residues: 2-31 <WAN>

R:Hayes, J.D.; Kerr, L.A.; Harrison, D.J.; Cronshaw, A.D.; Ross, A.G.; Neal, G.E.

Biochem. J. 268, 295-302, 1990

A:Title: Preferential over-expression of the class alpha rat Ya(2) glutathione S-trans-

ferase 1 (Ya1) and Ya(2) with cloned class alpha glutathione S-transferase cDNA sequences.

A:Reference number: S10555; MUID:90303203; PMID:2363675

A:Accession: S10555

A:Molecule type: protein

A:Residues: 17-50; 52-62; 64-85; 106-140; 209-222 <HAY>

R:Taylor, J.B.; Craig, R.K.; Beale, D.; Ketterer, B.

Biochem. J. 219, 223-231, 1984

A:Title: Construction and characterization of a plasmid containing complementary DNA to

A:Reference number: I52350; MUID:84202896; PMID:6547043

A:Accession: I52350

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-129 <REM>

A:Cross-references: EMBL:X00520; NID:G56329; PIRN:CAA25203.1; PID:G56330

R:Tu, C.P.D.; Weiss, M.J.; Karakawa, W.W.; Reddy, C.C.

Nucleic Acids Res. 10, 5407-5419, 1982

A:Title: Cloning and sequence analysis of a cDNA plasmid for one of the rat liver glut

A:Reference number: A93438; MUID:80064489; PMID:6292839

A:Accession: A93438

A:Molecule type: mRNA

A:Residues: 81-222 <TUC>

A:Experimental source: clone pGTR112

C:Comment: Glutathione transferases are multifunctional dimers of identical or similar

m-independent glutathione peroxidase activity toward organic hydroperoxides. In additi

on, and azocarcinogen dyes. It is a cytosolic protein found in many mammalian tissues

C:Function:

A:Description: catalyzes the conjugation of glutathione to electrophilic xenobiotics;

substrate hydrophobic ligands

C:Superfamily: glutathione transferase

C:Keywords: dimer; liver; transferase

F:2-222/Product: glutathione transferase class alpha chain Ya2 #status experimental <M

Query Match 35.0%; Score 53.5; DB 2; Length 222;

Best Local Similarity 53.8%; Pred. No. 6.9;

Matches 14; Conservative 3; Mismatches 8; Indels 1; Gaps 1;





V.  
 Nature 392, 353-358, 1998  
 A>Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A/Reference number: A70300; PMID:9819666; PMID:9537320  
 A/Accession: E70330  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-217 <AOF>  
 A/Cross-references: UNIPROT:O6670; GB:AE000684; NID:G2983009; PIDN:AA06632.1; PID:G298  
 A/Experimental source: strain VFS  
 C/Genetics:  
 A/Gene: aq338  
 C/Superfamily: spore germination protein C2; bioc homology  
 F/45-140/Domain: bioc homology <BiOC>

Query Match 34.0%; Score 52; DB 2; Length 217;  
 Best Local Similarity 42.9%; Pred. No. 11;  
 Matches 9; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

OY 3 MLYVQTEKYDLAIKDLKEAL 23  
 DB 162 GLLFGRDKWEFTKSLKESL 182

RESULT 14

T14686  
 hypothetical protein - Yersinia pestis plasmid pMT1  
 C/Species: Yersinia pestis  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: T14686  
 R/Hu, P.; Elliott, J.; McCreedy, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Carrano,  
 submitted to the EMBL Data Library, March 1998  
 A/Description: Structural organization of virulence determinants in three Yersinia pesti  
 A/Reference number: Z18168  
 A/Accession: T14686  
 A/Status: preliminary; translated from GB/EMBL/DDBJ  
 A/Molecule type: DNA  
 A/Residues: 1-317 <HUP>  
 A/Cross-references: UNIPROT:O68757; EMBL:AF053947; NID:G2996286; PID:G2996321; PIDN:AA01  
 C/Genetics:  
 A/Genome: plasmid pMT1

Query Match 34.0%; Score 52; DB 2; Length 317;  
 Best Local Similarity 37.0%; Pred. No. 16;  
 Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

OY 4 MLYVQTEKYDLAIKDLKEAL 30  
 DB 26 VLHPTETYEWIIKEIGLEVMLAGNH 52

RESULT 15

T14923  
 hypothetical protein Y1006 - Yersinia pestis plasmid pMT1  
 C/Species: Yersinia pestis  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: T14923  
 R/Lindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.  
 Infect. Immun. 66, 5731-5742, 1998  
 A>Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIMS plasmid  
 A/Reference number: Z18268; PMID:9943898; PMID:9826348  
 A/Accession: T14923  
 A/Status: preliminary; translated from GB/EMBL/DDBJ  
 A/Molecule type: DNA  
 A/Residues: 1-361 <LIN>  
 A/Cross-references: UNIPROT:Q9R2P6; EMBL:AF074611; NID:G3883003; PID:G3883008; PIDN:AA06  
 C/Genetics:  
 A/Gene: Y1006  
 A/Genome: plasmid pMT1

Query Match 34.0%; Score 52; DB 2; Length 361;  
 Best Local Similarity 37.0%; Pred. No. 19;  
 Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

OY 4 MLYVQTEKYDLAIKDLKEAL 30  
 DB 70 VLHPTETYEWIIKEIGLEVMLAGNH 96

Search completed: November 10, 2004, 12:29:04  
 Job time : 8.06481 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:27:01 ; Search time 32.8241 Seconds

(without alignments)  
525.871 Million cell updates/sec

Title: US-10-092-750-5

Perfect score: 153  
Sequence: 1 QRGMLXYQTEKYDLAKDKLXNALIQLRGN 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	96.1	526	1	NCF2_HUMAN
2	136	88.9	451	2	Q76EV7
3	136	88.9	451	2	BAD12546
4	132	86.3	525	2	Q70145
5	130	85.0	527	1	NCF2_BOVIN
6	130	85.0	527	2	Q95L70
7	128	83.7	526	2	Q9N0E9
8	119	77.8	543	2	Q95MN2
9	76	49.7	358	2	Q76EV9
10	76	49.7	358	2	BAD11767
11	74	48.4	534	2	Q6GMC8
12	72	47.1	201	2	Q9H311
13	66	43.1	254	2	Q74FV2
14	66	43.1	254	2	AAR33832
15	64	41.8	497	2	Q7PWF5
16	63.5	41.5	628	2	Q8RI47
17	63	41.2	241	2	Q8DGB3
18	62	40.5	174	2	Q8YRL6
19	62	40.5	436	2	Q8HUR6
20	62	40.5	444	2	Q8GSY2
21	62	40.5	444	2	Q8CU00
22	62	40.5	495	2	Q7U043
23	62	40.5	1321	1	IF3X_DICDI
24	60.5	39.5	175	2	Q7PWF4
25	60	39.2	911	2	Q7PWF2
26	60	39.2	1050	2	Q7JMW7
27	60	39.2	1050	2	AAS45307
28	58.5	38.2	320	2	P74450
29	58.5	38.2	369	2	Q7JL42
30	58.5	38.2	369	2	AAS12537
31	58	37.9	240	2	Q97CQ8

32	57	37.3	211	2	Q6VTR4	Q6VTR4 uncultured
33	57	37.3	211	2	Q6VUX6	Q6VUX6 uncultured
34	57	37.3	211	2	AAQ17285	AAQ17285 uncultured
35	57	37.3	211	2	AAQ17347	AAQ17347 uncultured
36	57	37.3	580	2	Q6DJ34	Q6DJ34 xenopus tro
37	56	36.6	221	1	GTAL_ANRST	GTAL_ANRST antechinus
38	56	36.6	520	1	TTCS_HUMAN	TTCS_HUMAN antechinus
39	56	36.6	520	2	Q89HE3	Q89HE3 homo sapien
40	56	36.6	593	2	Q8A7P1	Q8A7P1 bractynizob
41	55	35.9	173	2	Q7NNX2	Q7NNX2 bacteroides
42	55	35.9	230	2	Q8IA87	Q8IA87 gloeobacter
43	55	35.9	413	2	Q8IA88	Q8IA88 caenorhabdi
44	55	35.9	547	2	Q8Y265	Q8Y265 caenorhabdi
45	55	35.9	977	2	Q9M8Y0	Q9M8Y0 anabena sp
						Q9M8Y0 arabidopsis

#### ALIGNMENTS

RESULT 1	NCIF2_HUMAN	STANDARD;	PRT;	526 AA.
ID	NCIF2_HUMAN	STANDARD;	PRT;	526 AA.
AC	P19878; Q8NFC7; Q9BVS1;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Neutrophil cytosol factor 2 (NCF-2) (Neutrophil NADPH oxidase factor 2) (67 kDa neutrophil oxidase factor) (p67-phox) (NOXA2).			
GN	Name=NCF2;			
OC	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
CC	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90239566; PubMed=1692159;			
RA	Lieto T.L., Lomax K.J., Volpp B.D., Nunoi H., Sechler J.M.G.,			
RA	Nauseef W.M., Clark R.A., Gallin J.I., Malech H.L.;			
RT	"Cloning of a 67-kD neutrophil oxidase factor with similarity to a			
RT	noncatalytic region of p60c-src.";			
RL	Science 248:727-730(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94083655; PubMed=7903171;			
RA	Kenney R.T., Malech H.L., Epstein N.D., Roberts R.L., Lieto T.L.;			
RT	"Characterization of the p67phox gene: genomic organization and			
RT	restriction fragment length polymorphism analysis for prenatal			
RT	diagnosis in chronic granulomatous disease.";			
RL	Blood 82:3739-3740(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANTS ARG-181 AND LYS-328.			
RC	TISSUE=Colon adenocarcinoma;			
RX	PubMed=1207919; DOI=10.1016/S0006-291X(02)02059-4;			
RA	Yoshida I.S., Nishida S., Shimoyama T., Kawahara T., Rokutan K.,			
RA	Tsunawaki S.;			
RT	"Expression of a p67(phox) homolog in Caco-2 cells giving O(2)(-)-			
RT	reconstituting ability to cytochrome b(558) together with recombinant			
RT	p47(phox).";			
RL	Biochem. Biophys. Res. Commun. 296:1322-1328(2002).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANT ARG-181.			
RA	Kainine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,			
RA	Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,			
RA	Phehan M., Farmer A.;			
RT	"Cloning of human full-length cDNAs in BD Creator(TM) system donor			
RT	vector.";			
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	Lairol S.;			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A., AND VARIANT ARG-181.			

CC TISSUE=Lymphoma; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 CC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 CC Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 CC Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 CC Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 CC Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 CC Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 CC Stapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 CC Brownstein M.J., Ustin T.B., Toshynski S., Carninci P., Prange C.,  
 CC Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 CC Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 CC Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Halys S.W.,  
 CC Villalón D.K., Murty D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 CC Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 CC Whiting M., Madan A., Young A.C., Shenchenko Y., Bouffard G.G.,  
 CC Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 CC Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 CC Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 CC Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 CC "Generation and initial analysis of more than 15,000 full-length human  
 CC and mouse cDNA sequences";  
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC [7]  
 CC INTERACTION WITH SYTL1.  
 CC MEDLINE=21276368; PubMed=11278853; DOI=10.1074/jbc.M01167200;  
 CC McAdara Berkowitz J.K., Catz S.D., Johnson J.L., Ruedi J.M., Thon V.,  
 CC Bahior B.M.;  
 CC "UFRC, a novel tandem C2 domain-containing protein associated with the  
 CC leukocyte NADPH oxidase";  
 CC J. Biol. Chem. 276:18855-18862(2001).  
 CC [8]  
 CC X-RAY CRYSTALLOGRAPHY (2.4 ÅNGSTRÖMS) OF 1-203 IN COMPLEX WITH RAC1.  
 CC MEDLINE=2100498; PubMed=11090627;  
 CC Lapouge K., Smith S.J., Walker P.A., Gamblin S.J., Smerdon S.J.,  
 CC Rittinger K.;  
 CC "Structure of the TPR domain of p67phox in complex with Rac.GTP";  
 CC Mol. Cell 6:899-907(2000).  
 CC [9]  
 CC VARIANT AR-CGD GLUT-78.  
 CC MEDLINE=94114935; PubMed=8286749;  
 CC de Boer M., Hilarius-Stokman P.M., Hoesle J.P., Verhoeven A.J.,  
 CC Graf N., Kenney R.T., Seger R., Roos D.;  
 CC "Autosomal recessive chronic granulomatous disease with absence of the  
 CC 67-kD cytosolic NADPH oxidase component: identification of mutation  
 CC and detection of carriers";  
 CC Blood 83:531-536(1994).  
 CC [10]  
 CC VARIANT AR-CGD 160-GLU-VAL-161.  
 CC PubMed=9070911; DOI=10.1006/birc.1997.6204;  
 CC Bonizzato A., Russo M.P., Domin M., Dusi S.;  
 CC "Identification of a double mutation (Dis0V-K161E) (sic) in the  
 CC p67phox gene of a chronic granulomatous disease patient";  
 CC J. Biochem. Biophys. Res. Commun. 231:861-863(1997).  
 CC [11]  
 CC VARIANT AR-CGD 19-LYS-ASP-21 DEL.  
 CC PubMed=10498624;  
 CC Paliou F.J., Rae J., Noack D., Erickson R., Ding J.,  
 CC Garcia de Olatte D., Curmutte J.T.;  
 CC "Molecular characterization of autosomal recessive chronic  
 CC granulomatous disease caused by a defect of the nicotinamide adenine  
 CC dinucleotide phosphate (reduced form) oxidase component p67-phox";  
 CC J. Biol. Chem. 274:2505-2514(1999).  
 CC [12]  
 CC VARIANTS AR-CGD GLN-77 AND VAL-128.  
 CC PubMed=10538813;  
 CC Noack D., Rae J., Cross A.R., Munoz J., Salmen S., Mendoza J.A.,  
 CC Rossi N., Curmutte J.T., Heyworth P.G.;  
 CC "Autosomal recessive chronic granulomatous disease caused by novel  
 CC mutations in NCF2, the gene encoding the p67-phox component of  
 CC phagocyte NADPH oxidase";  
 CC Hum. Genet. 105:460-467(1999).  
 CC -1- FUNCTION: NCF2, NCF1, and a membrane bound cytochrome b558 are  
 CC required for activation of the latent NADPH oxidase (necessary for

CC superoxide production).  
 CC -1- SUBUNIT: Interacts with SYTL1 and RAC1.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- DISEASE: Defects in NCF2 are a cause of autosomal recessive  
 CC chronic granulomatous disease (AR-CGD) [MIM:233710]; also known as  
 CC autosomal cytochrome-b-positive chronic granulomatous disease. It  
 CC causes recurrent infection by catalase-positive organisms.  
 CC -1- SIMILARITY: Contains 2 SH3 domains.  
 CC -1- SIMILARITY: Contains 3 TPR repeats.  
 CC -1- DISEASE: NAME=NCF2base; NOTE=NCF2 deficiency database;  
 CC WWW=http://www.utu.fi/tmc/bioinfo/NCF2base/".  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M32011; AAA6379.1; -;  
 CC EMBL: U00788; AAB60320.1; -;  
 CC EMBL: U00776; AAB60320.1; JOINED.  
 CC EMBL: U00777; AAB60320.1; JOINED.  
 CC EMBL: U00778; AAB60320.1; JOINED.  
 CC EMBL: U00779; AAB60320.1; JOINED.  
 CC EMBL: U00780; AAB60320.1; JOINED.  
 CC EMBL: U00781; AAB60320.1; JOINED.  
 CC EMBL: U00782; AAB60320.1; JOINED.  
 CC EMBL: U00783; AAB60320.1; JOINED.  
 CC EMBL: U00784; AAB60320.1; JOINED.  
 CC EMBL: U00785; AAB60320.1; JOINED.  
 CC EMBL: U00786; AAB60320.1; JOINED.  
 CC EMBL: U00787; AAB60320.1; JOINED.  
 CC EMBL: AF527950; AAM89263.1; -;  
 CC EMBL: AF007439; AAP36107.1; -;  
 CC EMBL: AL137800; CAC19686.1; -;  
 CC EMBL: BC001606; AAH01606.1; -;  
 CC EMBL: P1R; A34855; A34855.  
 CC PDB: 1E96; X-ray; B=1-203.  
 CC PDB: 1H8; X-ray; A=1-213.  
 CC PDB: 1K4; NMR; S=455-516.  
 CC PDB: 1OEY; X-ray; A/B/C/D=347-429.  
 CC Genew; HSCN:7661; NCF2.  
 CC MIM: 608515; -;  
 CC GO: GO:0005489; F:electron transporter activity; TAS.  
 CC GO: GO:0006968; P:cellular defense response; TAS.  
 CC InterPro: IPR000108; Neu\_cyt\_fact\_2.  
 CC InterPro: IPR000270; OPR\_PBl.  
 CC InterPro: IPR001452; SH3.  
 CC InterPro: IPR001440; TPR.  
 CC InterPro: IPR008941; TPR-like.  
 CC Pfam: PF00564; PBl; 1.  
 CC Pfam: PF00018; SH3; 2.  
 CC Pfam: PF00515; TPR; 3.  
 CC PRINTS: PR00499; p67PHOX.  
 CC PRINTS: PR00452; SH3DOMAIN.  
 CC PRODOM: PD000066; SH3; 2.  
 CC PROSITE: PSS00002; SH3; 2.  
 CC PROSITE: PSS0005; TPR; 3.  
 CC PROSITE: PSS0293; TPR\_REGION; 1.  
 CC 3D-structure: Chronic granulomatous disease; Disease mutation;  
 CC Polymorphism: Repeat; SH3 domain; TPR repeat.  
 CC REPEAT 37 70 TPR 1.  
 CC REPEAT 71 104 TPR 2.  
 CC REPEAT 121 154 TPR 3.  
 CC DOMAIN 240 299 SH3 1.  
 CC DOMAIN 457 516 SH3 2.  
 CC VARIANT 19 21 Missing (in AR-CGD).  
 CC /FTId=VAR\_017387.  
 CC R -> O (in AR-CGD).  
 CC /FTId=VAR\_017388.

FT VARIANT 78 78 G-> E (in AR-CGD) /FTID=VAR\_008904.

Query Match  
Best Local Similarity 96.1%; Score 147; DB 1; Length 526;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORGMYYQTEKYDLAIKDLKALITQLRGN 29  
Db 76 ORGMYYQTEKYDLAIKDLKALITQLRGN 104

## RESULT 2

Q76EV7 PRELIMINARY; PRT; 451 AA.

AC Q76EV7  
DT 05-JUL-2004 (TREMblrel. 27, Created)  
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)  
DE Neutrophil cytosolic factor 2 (Fragment).  
GN Name=NCIF2;  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.  
OK NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=Hartley;  
RA Kawahara T., Sumimoto H., Kutsuni H., Kondo S.T., Kishi K.,  
Rokutan K.;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB105909; BAD12546.1; -  
DR InterPro; IPR000108; Neu\_cyt\_Fact\_2.  
DR InterPro; IPR000270; OPR\_P81.  
DR InterPro; IPR001452; SH3-  
DR InterPro; IPR001440; TPR.  
DR InterPro; IPR008941; TPR-like.  
DR Pfam; PF00564; P81; 1.  
DR Pfam; PF00018; SH3\_1; 2.  
DR Pfam; PF00515; TPR; 2.  
DR PRINTS; PR00499; P67PHOX.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR ProDom; PD000066; SH3; 2.  
DR SMART; SM00666; P81; 1.  
DR SMART; SM00326; SH3; 2.  
DR SMART; SM00028; TPR; 2.  
DR PROSITE; PS50002; SH3; 2.  
DR PROSITE; PS50005; TPR; 1.  
DR PROSITE; PS50293; TPR\_REGION; 1.  
KW Repeat; SH3 domain; TPR repeat.

FT NON\_TER 1  
FT NON\_TER 451  
SQ SEQUENCE 451 AA; 51664 MW; 70F2DCE8EDA60473 CRC64;

Query Match  
Best Local Similarity 93.1%; Score 136; DB 2; Length 451;  
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ORGMYYQTEKYDLAIKDLKALITQLRGN 29  
Db 33 ORGMYYQTEKYDLAIKDLKALITQLRGN 61

## RESULT 3

BAD12546 PRELIMINARY; PRT; 451 AA.

AC BAD12546  
DT 24-MAR-2004 (TREMblrel. 27, Created)  
DT 24-MAR-2004 (TREMblrel. 27, Last sequence update)  
DT 24-MAR-2004 (TREMblrel. 27, Last annotation update)  
DE Neutrophil cytosolic factor 2 (Fragment).  
GN NCIF2  
OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.  
OK NCBI\_TaxID=10141;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=Hartley;  
RA Kawahara T., Sumimoto H., Kutsuni H., Kondo S.T., Kishi K.,  
Rokutan K.;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB105909; BAD12546.1; -  
FT NON\_TER 1  
FT NON\_TER 451  
SQ SEQUENCE 451 AA; 51664 MW; 70F2DCE8EDA60473 CRC64;

Query Match  
Best Local Similarity 88.9%; Score 136; DB 2; Length 451;  
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ORGMYYQTEKYDLAIKDLKALITQLRGN 29

Db 33 ORGMYYQTEKYDLAIKDLKALITQLRGN 61

## RESULT 4

O70145 PRELIMINARY; PRT; 525 AA.

AC O70145  
DT 01-AUG-1998 (TREMblrel. 07, Created)  
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)  
DT 01-OCT-2004 (TREMblrel. 28, Last annotation update)  
DE p67phox (Neutrophil cytosolic factor 2) (Mus musculus adult male bone  
cDNA, RIKEN full-length enriched library, clone#1930001N11  
DE product:neutrophil cytosolic factor 2, full insert sequence).  
GN Name=NCIF2;  
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OK NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=Leukemia;

RX MEDLINE=98149672; PubMed=9490028;  
RA Mizuki K., Kadomatsu K., Hata K., Ito T., Fan Q.-W., Xage Y.,  
RA Fukumaki Y., Sakaki Y., Takeshige K., Sumimoto H.;  
RT "Functional modules and expression of mouse p40(phox) and p67(phox),  
RT SH3-domain-containing proteins involved in the phagocyte NADPH oxidase  
complex";  
RL Eur. J. Biochem. 251:573-582(1998).

RN [2]  
RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N;

RC TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old  
virgin mouse. Taken by biopsy;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang Y., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant P., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Baha S.S., Liguori N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez M.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Kravtsov M.I., Skalka U., Smalins D.E., Schnerch A., Schein J.E.,  
Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N;  
 RC TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old  
 RC virgin mouse. Taken by biopsy.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 RN [41]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RC MEDLINE=9279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RC MEDLINE=2108560; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RA The FANTOM Consortium;  
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RC MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RC MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system: 384-format sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Koide M., Koya S.,  
 RA Kuritara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shingawa K., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahita S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB002664; BAA2650.1; -  
 DR EMBL: BC003730; AA03730.1; -  
 DR EMBL: AK036379; BAC29404.1; -

DR HSSP; P19878; 1HH8.  
 DR MGD; MGI:97284; NCF2.  
 DR InterPro; IPR000108; Neu\_cyt\_fact\_2.  
 DR InterPro; IPR000270; OPR\_PBI\_  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001440; TPR.  
 DR InterPro; IPR008941; TPR-like.  
 DR Pfam; PF00564; PBI\_1.  
 DR Pfam; PF00564; SH3\_1; 2.  
 DR Pfam; PF00515; TPR; 3.  
 DR PRINTS; PR00499; P67PHOX.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 2.  
 DR SMART; SM00666; PBI; 1.  
 DR SMART; SM00326; SH3; 2.  
 DR SMART; SM00028; TPR; 3.  
 DR PROSITE; PS50002; SH3; 2.  
 DR PROSITE; PS50005; TPR; 1.  
 DR PROSITE; PS50293; TPR\_REGION; 1.  
 DR Repeat; SH3 domain; TPR repeat.  
 DR SEQUENCE 525 AA; 59485 KM; 663DB652D790F76 CRC64;  
 SQ  
 Query Match 86.3%; Score 132; DB 2; Length 525;  
 Best Local Similarity 89.7%; Pred. No. 4.1e-10;  
 Matches 26; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ORGNLYVQTEKYDLATKDKELIOLRGN 29  
 Db 76 QRGMLYRMEKYDLATKDKELIOLRGN 104  
 ID NCF2\_BOVIN STANDARD; PRT; 527 AA.  
 AC 07775;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Neutrophil cytosol factor 2 (NCF-2) (Neutrophil NADPH oxidase factor  
 DE 2) (67 kDa neutrophil oxidase factor) (p67-phox).  
 OS Name=NCF2;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OC NCBI\_Taxid=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20112525; PubMed=10647999;  
 RA Burger P.L., Swain S.D., Clements M.K., Siemsen D.W., Davis A.R.,  
 RA Gaus K.A., Quinn M.T.;  
 RT "Cloning and expression of bovine p47-phox and p67-phox: comparison with the human and murine homologs.";  
 RL J. Leukoc. Biol. 67:63-72(2000).  
 CC -!- FUNCTION: NCF2, NCF1, and a membrane bound cytochrome b558 are required for activation of the latent NADPH oxidase (necessary for superoxide production).  
 CC -!- SUBUNIT: Interacts with SYTL1 and PAC1 (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Contains 2 SH3 domains.  
 CC -!- SIMILARITY: Contains 3 TPR repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF079303; AAC82463.1; -  
 DR HSSP; P19878; 1HH8.  
 DR InterPro; IPR000108; Neu\_cyt\_fact\_2.

DR InterPro: IPR000270; OPR\_PBI.  
 DR InterPro: IPR001452; SH3.  
 DR InterPro: IPR001440; TPR.  
 DR InterPro: IPR008941; TPR-like.  
 DR Pfam: PF00564; PBI.1.  
 DR Pfam: PF00018; SH3.2.  
 DR Pfam: PF00515; TPR.3.  
 DR PRINTS: PR00499; P67PHOX.  
 DR PRINTS: PR00452; SH3DOMAIN.  
 DR ProDom: PD000066; SH3.2.  
 DR SMART: SM00326; SH3.2.  
 DR SMART: SM00028; TPR.3.  
 DR PROSITE: PS50002; SH3.2.  
 DR PROSITE: PS50005; TPR.3.  
 DR PROSITE: PS50293; TPR\_REGION.2.  
 KM Repeat: SH3 domain; TPR repeat.  
 FT REPEAT 37 70 TPR 1.  
 FT REPEAT 71 104 TPR 2.  
 FT REPEAT 121 154 TPR 3.  
 FT DOMAIN 240 299 SH3 1.  
 FT DOMAIN 458 517 SH3 2.  
 SQ SEQUENCE 527 AA; 59663 MW; 430D33FELD66BD1 CRC64;

Query Match 85.0%; Score 130; DB 1; Length 527;  
 Best Local Similarity 89.7%; Pred. No. 8e-10;  
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ORGMYYQTEKYDLAIKDLKEALIQLRGN 29  
 Db 76 ORGMYYQMEKYDSAIKDLKEALIQLRGN 104

RESULT 6  
 ID Q95L70 PRELIMINARY; PRT; 527 AA.  
 AC Q95L70;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE NADPH oxidase cytosolic protein p67phox.  
 OS Bison bison (American bison).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bison.  
 OC NCBI\_TaxID=9901;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE:22211501; PubMed:12223206;  
 RA Gauss K.A., Bunger P.L., Siemsen D.W., Young C.J., Nelson-Overton L.,  
 RA Pigge J.R., Swain S.D., Quinn M.T.;  
 RL "Molecular analysis of the bison phagocyte NADPH oxidase: cloning and  
 RT sequencing of five NADPH oxidase cDNAs";  
 RT Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 133:1-12(2002).  
 DR EMBL: AF411139; AAL11889.1; -.  
 DR HSSP: P19878; 1HH8.  
 DR InterPro: IPR000108; Neu\_Cyt\_fact\_2.  
 DR InterPro: IPR000270; OPR\_PBI.  
 DR InterPro: IPR001452; SH3.  
 DR InterPro: IPR001440; TPR.  
 DR InterPro: IPR008941; TPR-like.  
 DR Pfam: PF00564; PBI.1.  
 DR Pfam: PF00018; SH3.1.2.  
 DR Pfam: PF00515; TPR.3.  
 DR PRINTS: PR00499; P67PHOX.  
 DR PRINTS: PR00452; SH3DOMAIN.  
 DR ProDom: PD000066; SH3.2.  
 DR SMART: SM00326; SH3.2.  
 DR SMART: SM00028; TPR.3.  
 DR PROSITE: PS50002; SH3.2.  
 DR PROSITE: PS50005; TPR.1.  
 DR PROSITE: PS50293; TPR\_REGION.1.  
 KM Repeat: SH3 domain; TPR repeat.  
 SQ SEQUENCE 527 AA; 59694 MW; 43707D7FBCA5CB84 CRC64;

Query Match 85.0%; Score 130; DB 2; Length 527;  
 Best Local Similarity 89.7%; Pred. No. 8e-10;  
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ORGMYYQTEKYDLAIKDLKEALIQLRGN 29  
 Db 76 ORGMYYQMEKYDSAIKDLKEALIQLRGN 104

RESULT 7  
 ID Q9NOE9 PRELIMINARY; PRT; 526 AA.  
 AC Q9NOE9;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE P67-phox.  
 GN Name=P67phox;  
 OS Tursiops truncatus (Atlantic bottle-nosed dolphin).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;  
 OC Tursiops.  
 OC NCBI\_TaxID=9739;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21109079; PubMed=11182145;  
 RA Inoue Y., Itoh T., Jimbo T., Sakai T., Ueda K., Imaichi-Omi S.;  
 RT "Molecular cloning and identification of bottle-nosed dolphin  
 RT p40(phox), p47(phox) and p67(phox).";  
 RL Vet. Immunol. Immunopathol. 78:21-33(2001).  
 DR EMBL: AB035593; BAA9543.1; -.  
 DR HSSP: P19878; 1HH8.  
 DR InterPro: IPR000108; Neu\_Cyt\_fact\_2.  
 DR InterPro: IPR000270; OPR\_PBI.  
 DR InterPro: IPR001452; SH3.  
 DR InterPro: IPR001440; TPR.  
 DR InterPro: IPR008941; TPR-like.  
 DR Pfam: PF00564; PBI.1.  
 DR Pfam: PF00018; SH3.1.2.  
 DR Pfam: PF00515; TPR.3.  
 DR PRINTS: PR00499; P67PHOX.  
 DR PRINTS: PR00452; SH3DOMAIN.  
 DR ProDom: PD000066; SH3.2.  
 DR SMART: SM00666; PBI.1.  
 DR SMART: SM00326; SH3.2.  
 DR SMART: SM00028; TPR.3.  
 DR PROSITE: PS50002; SH3.2.  
 DR PROSITE: PS50005; TPR.1.  
 DR PROSITE: PS50293; TPR\_REGION.1.  
 KM Repeat: SH3 domain; TPR repeat.  
 SQ SEQUENCE 526 AA; 59363 MW; DB1459B6510DD82A CRC64;

Query Match 83.7%; Score 128; DB 2; Length 526;  
 Best Local Similarity 86.2%; Pred. No. 1.6e-09;  
 Matches 25; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ORGMYYQTEKYDLAIKDLKEALIQLRGN 29  
 Db 76 ORGMYYQMEKYDSALIKDLKEALIQLRGN 104

RESULT 8  
 ID Q95MN2 PRELIMINARY; PRT; 543 AA.  
 AC Q95MN2;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE P67-phox.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OK NCB1\_TaxID=9986;  
 RN [1]  
 RX MEDLINE=21676683; PubMed=11818454;  
 RA Gausa K.A., Mascio P.L., Siemsen D.W., Nelson L.K., Bunger P.L.,  
 RA Pagano P.J., Quinn M.T.;  
 RT "Cloning and sequencing of rabbit leukocyte NADPH oxidase genes  
 RT reveals a unique p67(phox) homolog.";  
 RL J. Leukoc. Biol. 71:319-326(2002).  
 DR EMBL: AF232789; AAK60124.1; -  
 DR HSSP: P19878; 1H8.  
 DR InterPro: IPR000108; Neu\_cyt\_fact\_2.  
 DR InterPro: IPR000270; OPR\_PBI.  
 DR InterPro: IPR001452; SH3.  
 DR InterPro: IPR001440; TPR.  
 DR InterPro: IPR008941; TPR-like.  
 DR Pfam: PF00564; PBI; 1.  
 DR Pfam: PF00018; SH3 1; 2.  
 DR Pfam: PF00515; TPR; 3.  
 DR PRINTS: PRO0499; P67PHOX.  
 DR PRINTS: PRO0452; SH3DOMAIN.  
 DR ProDom: PD000066; SH3; 2.  
 DR SMART: SM00666; PBI; 1.  
 DR SMART: SM00326; SH3; 2.  
 DR SMART: SM00028; TPR; 3.  
 DR PROSITE: PS50002; SH3; 2.  
 DR PROSITE: PS50005; TPR; 1.  
 DR PROSITE: PS50293; TPR\_REGION; 1.  
 DR Repeat: SH3 domain; TPR repeat.  
 KW SEQUENCE 543 AA; 60303 MW; 03284378B3924F4 CRC64;

Query Match 77.8%; Score 119; DB 2; Length 543;  
 Best Local Similarity 79.3%; Pred. No. 3, 1e-08;  
 Matches 23; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ORGMATYOTEKYDIAIKDKKALIQLRGN 29  
 Db 76 QRGMLHRCRBYDAIKDKKALQLRGN 104

RESULT 9  
 Q76EV9 PRELIMINARY; PRT; 358 AA.  
 ID Q76EV9;  
 AC 076EV9;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE NADPH oxidase activator 1 (Fragment).  
 GN Name=NOXAL;  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.  
 OC NCB1\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hartley;  
 RA PubMed=14978110;  
 RA Kawahara T., Kuwano Y., Teshima-Kondo S., Takeya R., Sumimoto H.,  
 RA Kishi K., Tanunawaki S., Hirayama T., Rokutan K.;  
 RA "Role of nicotinamide adenine dinucleotide phosphate oxidase 1 in  
 RT oxidative burst response to toll-like receptor 5 signaling in large  
 RT intestinal epithelial cells.";  
 RL J. Immunol. 172:3051-3058(2004).  
 DR EMBL: AB105907; BAD11767.1; -  
 DR InterPro: IPR001452; SH3.  
 DR InterPro: IPR001440; TPR.  
 DR InterPro: IPR008941; TPR-like.  
 DR Pfam: PF00515; TPR; 2.  
 DR SMART: SM00028; TPR; 2.  
 DR NON\_TER 1  
 FT NON\_TER 358  
 SQ SEQUENCE 358 AA; 39033 MW; 171E02E98980D927 CRC64;

Query Match 49.7%; Score 76; DB 2; Length 358;  
 Best Local Similarity 43.3%; Pred. No. 0.029;  
 Matches 13; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ORGMATYOTEKYDIAIKDKKALIQLRGN 30  
 Db 7 QRGVAHFQDQFQALSDFQALVQLRGNS 36

RESULT 10  
 BAD11767 PRELIMINARY; PRT; 358 AA.  
 ID BAD11767;  
 AC BAD11767;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE NADPH oxidase activator 1 (Fragment).  
 GN NOXAL.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.  
 OC NCB1\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hartley;  
 RA Kawahara T., Kuwano Y., Teshima-Kondo S., Takeya R., Sumimoto H.,  
 RA Kishi K., Tanunawaki S., Hirayama T., Rokutan K.;  
 RA "Role of Nicotinamide Adenine Dinucleotide Phosphate Oxidase 1 in  
 RT oxidative burst response to Toll-like Receptor 5 signaling in large  
 RT intestinal epithelial cells.";  
 RL J. Immunol. 172:3051-3058(2004).  
 DR EMBL: AB105907; BAD11767.1; -  
 DR NON\_TER 1  
 FT NON\_TER 358  
 SQ SEQUENCE 358 AA; 39033 MW; 171E02E98980D927 CRC64;

Query Match 49.7%; Score 76; DB 2; Length 358;  
 Best Local Similarity 43.3%; Pred. No. 0.029;  
 Matches 13; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ORGMATYOTEKYDIAIKDKKALIQLRGN 30  
 Db 7 QRGVAHFQDQFQALSDFQALVQLRGNS 36

RESULT 11  
 Q6GMC8 PRELIMINARY; PRT; 534 AA.  
 ID Q6GMC8;  
 AC Q6GMC8;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCB1\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Siemsen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Mulvaney S.J.,  
 RA Rana S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mulvaney S.J.,  
 RA Bosak S.A., McGowan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulik S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,



```

RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16893-16903(2002).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN (3)
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RL Klein S., Strausberg R.,
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC074140; AAH74140.1; -
DR InterPro; IPR000108; Neu_cyt_fac2_2.
DR InterPro; IPR000270; OPR_PBI.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR011511; SH3_2.
DR InterPro; IPR001440; TPR_1like.
DR InterPro; IPR008941; TPR_1like.
DR Pfam; PF00564; PBI; 1.
DR Pfam; PF00018; SH3_1; 2.
DR Pfam; PF07653; SH3_2; 2.
DR Pfam; PF00515; TPR; 3.
DR PRINTS; PR00499; P67PHOX.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRODom; PD000066; SH3; 2.
DR SMART; SM00666; PBI; 1.
DR SMART; SM00326; SH3; 2.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS50002; SH3; 2.
DR PROSITE; PS50293; TPR_REGION; 1.
DR Hypothetical protein; Repeat; SH3 domain; TPR repeat.
KW SEQUENCE 534 AA; 60270 MW; 8B8D95618C05515 CRC64;
SQ
Query Match 48.4%; Score 74; DB 2; Length 534;
Best Local Similarity 48.3%; Pred. No. 0.087;
Matches 14; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
QY 1 QGMLYYQTEKYDLAIKDKLKEALIKRGK 29
DB 76 QRGYVYFQGRKYNLAQODMSRAVYTEMKRN 104

```

```

RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513(2000).
DR EMBL; AL445063; CAC1187.1; -.
DR HSSP; P31948; IEHW.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR_1like.
DR Pfam; PF00515; TPR; 5.
DR SMART; SM00028; TPR; 5.
DR PROSITE; PS50005; TPR; 3.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Complete proteome.
SQ SEQUENCE 201 AA; 23169 MW; 535B99EFF8EC81D4 CRC64;
QY Query Match 47.1%; Score 72; DB 2; Length 201;
Best Local Similarity 59.1%; Pred. No. 0.056;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
DB 21 QGMAYIMKKYKEBAIKDKKSI 42
RESULT 13
ID Q7AFV2 PRELIMINARY; PRT; 254 AA.
AC Q7AFV2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Lipoprotein, putative.
GN ORFNames=GSTU0501;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Heide B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Weidman J.F., Khoult H.M., Feldblyum T.V., Uitterlinden H.A.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA Daviden J.F., Khoult H.M., Feldblyum T.V., Uitterlinden H.A.,
RA Van Aken S.E., Lovley D.R., Fraser C.M.;
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments.";
RL Science 302:1967-1969(2003).
DR EMBL; AE017180; AAR3832.1; -.
DR TIGR; GSTU0501; -.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR_1like.
DR InterPro; IPR005156; UPF0169.
DR Pfam; PF01696; UPF0169; 1.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
DR Lipoprotein.
KW SEQUENCE 254 AA; 29398 MW; 9EF399C04666F582 CRC64;
QY Query Match 43.1%; Score 66; DB 2; Length 254;
Best Local Similarity 61.9%; Pred. No. 0.53;
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
DB 3 GMLYYQTEKYDLAIKDKLKEAL 23
DB 177 QGFYRTKXTYSAIKRLDAL 197
RESULT 14
AAR3832 PRELIMINARY; PRT; 254 AA.
ID AAR3832

```

AC AAR33832;  
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Lipoprotein, putative.  
 GN GSU0501.  
 OS Geobacter sulfurreducens.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;  
 CC Geobacteraceae; Geobacter.  
 OX NCBI\_TaxID=35554;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PC9 / ATCC 51573;  
 RX PubMed=14671304;  
 RA Mehta B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,  
 Heidelberg J.F., Wu D., Ward N.L., Beanan M.J., Dodson R.J.,  
 Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,  
 Rasmussen M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,  
 Davisson T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,  
 Weidman J., Khouri H.M., Feldblyum T.V., Utterback T.R.,  
 Van Aken S.E., Lovley D.R., Fraser C.M.,  
 "Genome of Geobacter sulfurreducens: metal reduction in subsurface  
 environments."  
 RT Science 302:1967-1969(2003).  
 DR EMBL: AS017208; AAR33832.1; -.  
 DR TIGR: GSU0501; -.  
 KM Lipoprotein  
 SQ SEQUENCE 254 AA; 29398 MW; 9EF399C0466F582 CRC64;

Query Match 43.1%; Score 66; DB 2; Length 254;  
 Best Local Similarity 61.9%; Pred. No. 0.53;  
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 GMLYYQTEKYDLAIKDLKEAL 23  
 DB 177 GQFYRYTEKYSIAIKRLDAL 197

## RESULT 15

Q7PWF5 PRELIMINARY; PRT; 497 AA.  
 ID Q7PWF5;  
 AC Q7PWF5;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE ENSANGP00000019419 (Fragment).  
 GN Name=ENSANGS00000016930;  
 OS Acropora gemmifera str. PEST.  
 OC Eukaryota; Metazoa; Anthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases;  
 CC -! CAUTION: The sequence shown here is derived from an  
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 preliminary data.  
 CC EMBL: AAB01008984; EAA14869.1; -.  
 DR InterPro: IPR001623; DnaJ\_N.  
 DR InterPro: IPR001440; TPR.  
 DR Pfam: PF00226; DnaJ\_1.  
 DR Pfam: PF00515; TPR; 8.  
 DR PROSITE: PS50076; DnaJ\_2; 1.  
 DR PROSITE: PS50005; TPR; 5.  
 DR PROSITE: PS50293; TPR\_REGION; 1.  
 FT NON TER 1  
 SQ SEQUENCE 497 AA; 56845 MW; 7AF3FBEBB98316A CRC64;

Query Match 41.8%; Score 64; DB 2; Length 497;  
 Best Local Similarity 52.2%; Pred. No. 2.2;  
 Matches 12; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 QRGMLYYQTEKYDLAIKDLKEAL 23  
 DB 332 QRAKLYNNMNEYEAAYDYKEL 354

Search completed: November 10, 2004, 12:26:48  
 Job time : 35.8241 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 ; Search time 11.3179 Seconds  
(without alignments)

222.664 Million cell updates/sec

Title: US-10-092-750-6  
Perfect score: 181  
Sequence: 1 GSESDPHFDALMQLAKAVASAAALVLAKSVAGR 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/iaa/5A\_COMB pep:\*
- 2: /cgn2\_6/prodata/1/iaa/5B\_COMB pep:\*
- 3: /cgn2\_6/prodata/1/iaa/6A\_COMB pep:\*
- 4: /cgn2\_6/prodata/1/iaa/6B\_COMB pep:\*
- 5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB pep:\*
- 6: /cgn2\_6/prodata/1/iaa/backfillseq1 pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	175	96.7	1024	US-09-562-737-41	Sequence 41, Appl
2	157	86.7	1024	US-09-562-737-43	Sequence 43, Appl
3	157	86.7	1024	US-09-562-737-44	Sequence 44, Appl
4	157	86.7	1024	US-09-562-737-47	Sequence 47, Appl
5	156	86.2	1024	US-09-562-737-48	Sequence 48, Appl
6	154	85.1	1024	US-09-562-737-45	Sequence 45, Appl
7	151	83.4	1024	US-09-562-737-50	Sequence 50, Appl
8	150	82.9	1024	US-09-562-737-49	Sequence 49, Appl
9	149	82.3	1024	US-09-562-737-42	Sequence 42, Appl
10	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
11	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
12	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
13	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
14	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
15	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
16	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
17	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
18	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
19	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
20	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
21	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
22	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
23	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
24	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
25	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
26	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl
27	147	81.2	1024	US-09-562-737-46	Sequence 46, Appl

28	49	27.1	189	3	US-09-201-945-74	Sequence 74, Appl
29	49	27.1	472	3	US-09-004-838-48	Sequence 48, Appl
30	49	27.1	472	3	US-09-004-838-103	Sequence 103, Appl
31	48.5	26.8	99	4	US-09-252-991A-22347	Sequence 22347, A
32	48.5	26.8	396	4	US-09-248-796A-20351	Sequence 20351, A
33	48	26.5	216	4	US-09-583-110-4709	Sequence 4709, Ap
34	48	26.5	272	4	US-09-252-991A-17682	Sequence 17682, A
35	48	26.5	550	4	US-09-252-991A-17682	Sequence 17682, A
36	48	26.5	588	4	US-09-489-039A-13579	Sequence 13579, A
37	48	26.5	815	4	US-09-583-110-4101	Sequence 4101, Ap
38	48	26.5	1147	1	US-08-121-365B-38	Sequence 38, Appl
39	48	26.5	1147	2	US-08-668-123-38	Sequence 38, Appl
40	47.5	26.2	407	4	US-09-583-110-3040	Sequence 3040, Ap
41	47.5	26.2	485	1	US-07-861-075-1	Sequence 1, Appl
42	47.5	26.2	485	1	US-08-120-827-1	Sequence 1, Appl
43	47.5	26.2	485	1	US-08-478-675-1	Sequence 1, Appl
44	47	26.0	446	4	US-09-489-039A-12283	Sequence 12283, A
45	47	26.0	478	3	US-09-004-838-105	Sequence 105, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-562-737-41
; Sequence 41, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: USW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-41

Query Match 96.7%; Score 175; DB 4; Length 1024;
Best Local Similarity 100.0%; Pred. No. 1.2e-17;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSESDPHFDALMQLAKAVASAAALVLAKSVAGR 38
DB 656 GSESDPHFDALMQLAKAVASAAALVLAKSVAGR 692

RESULT 2
US-09-562-737-43
; Sequence 43, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: USW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
```

```
/ OTHER INFORMATION: Sequence
US-09-562-737-43

Query Match
Best Local Similarity 86.7%; Score 157; DB 4; Length 1024;
Pred. No. 6.1e-15;
Matches 34; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 38
DB 656 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 692

RESULT 3
US-09-562-737-44
; Sequence 44, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 44
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-44

Query Match
Best Local Similarity 86.7%; Score 157; DB 4; Length 1024;
Pred. No. 6.1e-15;
Matches 34; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 38
DB 656 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 692

RESULT 4
US-09-562-737-47
; Sequence 47, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 47
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-47

Query Match
Best Local Similarity 86.7%; Score 157; DB 4; Length 1024;
Pred. No. 6.1e-15;
Matches 33; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 38
DB 656 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 692

OTHER INFORMATION: Sequence
US-09-562-737-43

Query Match
Best Local Similarity 86.7%; Score 157; DB 4; Length 1024;
Pred. No. 6.1e-15;
Matches 33; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 38
DB 656 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 692

RESULT 5
US-09-562-737-48
; Sequence 48, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 48
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-48

Query Match
Best Local Similarity 86.2%; Score 156; DB 4; Length 1024;
Pred. No. 8.6e-15;
Matches 33; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 38
DB 656 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 692

RESULT 6
US-09-562-737-45
; Sequence 45, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-45

Query Match
Best Local Similarity 85.1%; Score 154; DB 4; Length 1024;
Pred. No. 1.7e-14;
Matches 34; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 38
DB 656 GESDTPHFQDAMQAKAVASAAALVTKAKSVQR 692

RESULT 7
US-09-562-737-50
; Sequence 50, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
```

;/ CURRENT FILING DATE: 2000-05-01  
;/ NUMBER OF SEQ ID NOS: 132  
;/ SOFTWARE: Patentin Ver. 2.1  
;/ SEQ ID NO 50  
;/ LENGTH: 1024  
;/ TYPE: PRT  
;/ ORGANISM: Artificial Sequence  
;/ FEATURE:  
;/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-562-737-50

Query Match 83.4%; Score 151; DB 4; Length 1024;  
Best Local Similarity 89.2%; Pred. No. 4,9e-14;  
Matches 33; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 38  
DB 656 GESDVPHFQDALMQLAKAVASAAALVTKAKSVQR 692

RESULT 8  
US-09-562-737-49  
;/ Sequence 49, Application US/09562737  
;/ Patent No. 6428967  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Herz, Joachim  
;/ APPLICANT: Gotthardt, Michael  
;/ TITLE OF INVENTION: LDL Receptor Signaling Pathways  
;/ FILE REFERENCE: UTSW0708  
;/ CURRENT APPLICATION NUMBER: US/09/562,737  
;/ CURRENT FILING DATE: 2000-05-01  
;/ NUMBER OF SEQ ID NOS: 132  
;/ SOFTWARE: Patentin Ver. 2.1  
;/ SEQ ID NO 49  
;/ LENGTH: 1024  
;/ TYPE: PRT  
;/ ORGANISM: Artificial Sequence  
;/ FEATURE:  
;/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-562-737-49

Query Match 82.9%; Score 150; DB 4; Length 1024;  
Best Local Similarity 89.2%; Pred. No. 6,9e-14;  
Matches 33; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 38  
DB 656 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 692

RESULT 9  
US-09-562-737-42  
;/ Sequence 42, Application US/09562737  
;/ Patent No. 6428967  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Herz, Joachim  
;/ APPLICANT: Gotthardt, Michael  
;/ TITLE OF INVENTION: LDL Receptor Signaling Pathways  
;/ FILE REFERENCE: UTSW0708  
;/ CURRENT APPLICATION NUMBER: US/09/562,737  
;/ CURRENT FILING DATE: 2000-05-01  
;/ NUMBER OF SEQ ID NOS: 132  
;/ SOFTWARE: Patentin Ver. 2.1  
;/ SEQ ID NO 42  
;/ LENGTH: 1024  
;/ TYPE: PRT  
;/ ORGANISM: Artificial Sequence  
;/ FEATURE:  
;/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-562-737-42

Query Match 82.3%; Score 149; DB 4; Length 1024;  
Best Local Similarity 91.7%; Pred. No. 9,8e-14;  
Matches 33; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 37  
DB 656 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 691

RESULT 10  
US-09-562-737-46  
;/ Sequence 46, Application US/09562737  
;/ Patent No. 6428967  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Herz, Joachim  
;/ APPLICANT: Gotthardt, Michael  
;/ TITLE OF INVENTION: LDL Receptor Signaling Pathways  
;/ FILE REFERENCE: UTSW0708  
;/ CURRENT APPLICATION NUMBER: US/09/562,737  
;/ CURRENT FILING DATE: 2000-05-01  
;/ NUMBER OF SEQ ID NOS: 132  
;/ SOFTWARE: Patentin Ver. 2.1  
;/ SEQ ID NO 46  
;/ LENGTH: 1024  
;/ TYPE: PRT  
;/ ORGANISM: Artificial Sequence  
;/ FEATURE:  
;/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-562-737-46

Query Match 81.2%; Score 147; DB 4; Length 1024;  
Best Local Similarity 91.7%; Pred. No. 2e-13;  
Matches 33; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 ESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 38  
DB 657 ESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 692

RESULT 11  
US-09-252-991A-19345  
;/ Sequence 19345, Application US/09252991A  
;/ Patent No. 6551795  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Marc J. Rubenfield et al.  
;/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
;/ FILE REFERENCE: 107196.136  
;/ CURRENT APPLICATION NUMBER: US/09/252,991A  
;/ CURRENT FILING DATE: 1999-02-18  
;/ PRIOR APPLICATION NUMBER: US 60/074,788  
;/ PRIOR FILING DATE: 1998-02-18  
;/ PRIOR APPLICATION NUMBER: US 60/094,190  
;/ PRIOR FILING DATE: 1998-07-27  
;/ NUMBER OF SEQ ID NOS: 33142  
;/ SEQ ID NO 19345  
;/ LENGTH: 283  
;/ TYPE: PRT  
;/ ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19345

Query Match 30.7%; Score 55.5; DB 4; Length 283;  
Best Local Similarity 41.9%; Pred. No. 2.5;  
Matches 18; Conservative 4; Mismatches 14; Indels 7; Gaps 2;

OY 2 GESDTPHFQD-----ALMQLAKAVASAAALVTKAKSVQR 38  
DB 12 GELVADP-LQDRQRRLQTVAGVAVALLLVLAQAQAVOR 53

RESULT 12

```

US-09-270-767-59294
; Sequence 59294, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 59294
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-59294

Query Match
Best Local Similarity 46.2%; Score 54; DB 4; Length 127;
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 6 TDHFDQALMQLAKAVASAAALVLK 31
DB 52 TSPOFOALAFSSALQSLQGLPVIX 77

RESULT 13
US-09-270-767-43890
; Sequence 43890, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 43890
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-43890

Query Match
Best Local Similarity 46.2%; Score 54; DB 4; Length 384;
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 6 TDHFDQALMQLAKAVASAAALVLK 31
DB 309 TSPOFOALAFSSALQSLQGLPVIX 334

RESULT 14
US-09-489-039A-8011
; Sequence 8011, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8011
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8011

```

```

Query Match
Best Local Similarity 29.3%; Score 53; DB 4; Length 562;
Matches 12; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 3 ESDTDPHFDQALMQLAKAVASAAALVLKXKS 34
DB 179 QGDLDPHFSTVADENAFPOAEVVAQALQMLAQS 210

RESULT 15
US-09-248-796A-24946
; Sequence 24946, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24946
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24946

Query Match
Best Local Similarity 28.2%; Score 51; DB 4; Length 74;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGESDTPDPHF 10
DB 48 GGEADNDPDPHF 57

```

Search completed: November 10, 2004, 12:32:21  
Job time: 12.3179 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 ; Search time 36.6512 Seconds  
(without alignments)  
366.225 Million cell updates/sec

Title: US-10-092-750-6  
Perfect score: 181  
Sequence: 1 GGSBDTPHFQDALMQLAKAVASAAALVLRKASVQR 38

Scoring table: BLOSUM62

Searched: 1566620 seqs, 35322586 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09 NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10 NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11 NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60 NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	181	100.0	38	US-10-092-750-6
2	175	96.7	1024	US-10-211-962-41
3	175	96.7	2541	US-10-177-293-470
4	157	86.7	1024	US-10-211-962-43
5	157	86.7	1024	US-10-211-962-44
6	157	86.7	1024	US-10-211-962-47
7	156	86.2	1024	US-10-211-962-48
8	154	85.1	1024	US-10-211-962-45
9	151	83.4	1024	US-10-211-962-50
10	150	82.9	1024	US-10-211-962-49
11	149	82.3	1024	US-10-211-962-42
12	147	81.2	1024	US-10-211-962-46
13	127	70.2	2545	US-10-092-900A-76

14	73.5	40.6	245	15	US-10-264-049-2981	Sequence 2981, Ap
15	73.5	40.6	698	14	US-10-043-487-309	Sequence 309, App
16	64.5	35.6	949	16	US-10-408-765A-268	Sequence 268, App
17	61.5	34.0	80	11	US-09-864-408A-8474	Sequence 8474, Ap
18	54	29.8	188	17	US-10-425-115-314807	Sequence 314807, A
19	54	29.8	326	15	US-10-424-599-176547	Sequence 176547, A
20	54	29.8	328	15	US-10-296-115-800	Sequence 800, App
21	54	29.8	653	10	US-09-374-046A-20	Sequence 20, App1
22	54	29.8	653	14	US-10-102-524-1851	Sequence 1851, Ap
23	54	29.8	653	14	US-10-331-496A-81	Sequence 81, App1
24	54	29.8	653	15	US-10-616-263-20	Sequence 20, App1
25	54	29.8	1073	14	US-10-156-761-12156	Sequence 12156, A
26	53	29.3	427	15	US-10-425-114-38212	Sequence 38212, A
27	53	29.3	548	15	US-10-282-122A-59635	Sequence 59635, A
28	53	29.3	1066	14	US-10-239-431A-28	Sequence 28, App1
29	53	29.3	1066	15	US-10-440-464-181	Sequence 181, App
30	52.5	29.0	173	16	US-10-437-963-203772	Sequence 203772, A
31	52	28.7	221	15	US-10-424-599-273855	Sequence 273855, A
32	52	28.7	249	14	US-10-080-170-363	Sequence 363, App
33	52	28.7	249	16	US-10-080-170-363	Sequence 363, App
34	52	28.7	249	17	US-10-468-356-363	Sequence 363, App
35	51.5	28.5	462	9	US-09-738-626-5888	Sequence 5888, Ap
36	51.5	28.5	462	16	US-10-781-014-478	Sequence 478, App
37	51	28.2	812	14	US-10-369-493-19403	Sequence 19403, A
38	51	28.2	3170	15	US-10-329-148A-4	Sequence 4, App1
39	50.5	27.9	569	9	US-09-738-626-6179	Sequence 6179, Ap
40	50.5	27.9	569	15	US-10-627-476-20	Sequence 20, App1
41	50.5	27.9	581	15	US-10-282-122A-65905	Sequence 65905, A
42	50.5	27.9	787	15	US-10-282-122A-72418	Sequence 72418, A
43	50	27.6	30	16	US-10-667-004-1	Sequence 1, App1
44	50	27.6	106	17	US-10-425-115-357579	Sequence 357579, A
45	49.5	27.3	517	15	US-10-072-012-437	Sequence 437, App

## ALIGNMENTS

RESULT 1  
US-10-092-750-6, Application US/10092750  
; Sequence 6, Application US/10092750  
; Publication No. US20030032157A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; APPLICANT: Alpin, Julia  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL  
; FILE REFERENCE: 50036/050002  
; CURRENT APPLICATION NUMBER: US/10/092,750  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: US 60/274,526  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-750-6

Query Match 100.0%; Score 181; DB 14; Length 38;  
Best Local Similarity 100.0%; Pred. No. 2.4e-18;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGSBDTPHFQDALMQLAKAVASAAALVLRKASVQR 38  
DB 1 GGSBDTPHFQDALMQLAKAVASAAALVLRKASVQR 38

RESULT 2  
US-10-211-962-41  
; Sequence 41, Application US/10211962  
; Publication No. US20030082640A1  
; GENERAL INFORMATION:

```
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-10-211-962-41

Query Match          96.7%; Score 175; DB 14; Length 1024;
Best Local Similarity 100.0%; Pred. No. 9.1e-15;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 38
DB 656 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 692

RESULT 3
US-10-177-293-470
; Sequence 470, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Weric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 470
; LENGTH: 2541
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; US-10-177-293-470

Query Match          96.7%; Score 175; DB 14; Length 2541;
Best Local Similarity 100.0%; Pred. No. 2.7e-15;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 38
DB 656 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 692

RESULT 4
US-10-211-962-43
; Sequence 43, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-10-211-962-43

Query Match          86.7%; Score 157; DB 14; Length 1024;
Best Local Similarity 91.9%; Pred. No. 3.3e-13;
Matches 34; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 38
DB 656 GESDTPHFQDALMQLAKAVASAAALVTKAKSVQR 692

RESULT 5
US-10-211-962-44
; Sequence 44, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-10-211-962-44

Query Match          86.7%; Score 157; DB 14; Length 1024;
Best Local Similarity 91.9%; Pred. No. 3.3e-13;
Matches 34; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```



QY 2 GESDTPHFQDALMQLAKAVASAAAALVTKAKSVQR 38  
DB 656 GESDTPHFQDALMQLAKAVASAAAALVTKAKSVQR 692

RESULT 6  
US-10-211-962-47  
; Sequence 47, Application US/10211962  
; Publication No. US20030082640A1  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/10/211,962  
; CURRENT FILING DATE: 2002-08-01  
; PRIOR APPLICATION NUMBER: US/09/562,737  
; PRIOR FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 47  
; LENGTH: 1024  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; US-10-211-962-47

Query Match 86.7%; Score 157; DB 14; Length 1024;  
Best Local Similarity 89.2%; Pred. No. 3,3e-13;  
Matches 33; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GESDTPHFQDALMQLAKAVASAAAALVTKAKSVQR 38  
DB 656 GRSDDPHFQDSLWQLAKAVASAAAALVTKAKSVQR 692

RESULT 7  
US-10-211-962-48  
; Sequence 48, Application US/10211962  
; Publication No. US20030082640A1  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/10/211,962  
; CURRENT FILING DATE: 2002-08-01  
; PRIOR APPLICATION NUMBER: US/09/562,737  
; PRIOR FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 48  
; LENGTH: 1024  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; US-10-211-962-48

Query Match 86.2%; Score 156; DB 14; Length 1024;  
Best Local Similarity 89.2%; Pred. No. 4,6e-13;  
Matches 33; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GESDTPHFQDALMQLAKAVASAAAALVTKAKSVQR 38  
DB 656 GETDTPHFQDALMQLAKAVASAAAALVTKAKSVQR 692

RESULT 8  
US-10-211-962-45

; Sequence 45, Application US/10211962  
; Publication No. US20030082640A1  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/10/211,962  
; CURRENT FILING DATE: 2002-08-01  
; PRIOR APPLICATION NUMBER: US/09/562,737  
; PRIOR FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 45  
; LENGTH: 1024  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; US-10-211-962-45

Query Match 85.1%; Score 154; DB 14; Length 1024;  
Best Local Similarity 91.9%; Pred. No. 8,8e-13;  
Matches 34; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GESDTPHFQDALMQLAKAVASAAAALVTKAKSVQR 38  
DB 656 GESDTPHFQDALMQLAKAVASAAAALVTKAKSVQR 692

RESULT 9  
US-10-211-962-50  
; Sequence 50, Application US/10211962  
; Publication No. US20030082640A1  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/10/211,962  
; CURRENT FILING DATE: 2002-08-01  
; PRIOR APPLICATION NUMBER: US/09/562,737  
; PRIOR FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 50  
; LENGTH: 1024  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; US-10-211-962-50

Query Match 83.4%; Score 151; DB 14; Length 1024;  
Best Local Similarity 89.2%; Pred. No. 2,4e-12;  
Matches 33; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GESDTPHFQDALMQLAKAVASAAAALVTKAKSVQR 38  
DB 656 GESDVPHFQDALMQLAKAVASAAAALVTKAKSVQR 692

RESULT 10  
US-10-211-962-49  
; Sequence 49, Application US/10211962  
; Publication No. US20030082640A1  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708

```

; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 49:
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-10-211-962-49

Query Match      82.9%; Score 150; DB 14; Length 1024;
Best Local Similarity 89.2%; Pred. No. 3.3e-12;
Matches 33; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 GESDTPHFQDALMQLAKAVASAAAALVLRKASVACR 38
DB      656 GESRDPHFQDALMSQLAKAVASATATLALVLRKASVACR 692

RESULT 11
; Sequence 42, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 42
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-10-211-962-42

Query Match      82.3%; Score 149; DB 14; Length 1024;
Best Local Similarity 91.7%; Pred. No. 4.5e-12;
Matches 33; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GESDTPHFQDALMQLAKAVASAAAALVLRKASVACR 37
DB      656 GESDTPHFQDALMQLAKAVASAAAALVLRKASVACR 691

RESULT 12
; Sequence 46, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 46

; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-10-092-900A-76

Query Match      81.2%; Score 147; DB 14; Length 1024;
Best Local Similarity 91.7%; Pred. No. 8.7e-12;
Matches 33; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3 ESDTDPHFQDALMQLAKAVASAAAALVLRKASVACR 38
DB      657 ESDTDPHFQDALMQLAKAVASAAAALVLRKASVACR 692

RESULT 13
; Sequence 76, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muraidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zethusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patrudajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vermet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malysankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderma, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Hainong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/336,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,261
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
```

PRIOR APPLICATION NUMBER: USSN 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: USSN 60/287,424  
PRIOR FILING DATE: 2001-04-30  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 768  
SEQ ID NO 76  
LENGTH: 2545  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-092-900A-76

Query Match 70.2%; Score 127; DB 15; Length 2545;  
Best Local Similarity 75.0%; Pred. No. 1,8e-08;  
Matches 27; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 GSDTDPHFQDALMQLAKAVASAAALVTKAKSVAQ 37  
DB 662 GENETDERFQDVLMSLAKAVANAAMLVTKAKNVAQ 697

RESULT 14  
US-10-264-049-2981  
Sequence 2981, Application US/10264049  
Publication No. US20040005579A1  
GENERAL INFORMATION:  
APPLICANT: Birse et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA133PI  
CURRENT APPLICATION NUMBER: US/10/264,049  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/18569  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: US 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: PatentIn Ver. 3.1  
SEQ ID NO 2981  
LENGTH: 245  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-264-049-2981

Query Match 40.6%; Score 73.5; DB 15; Length 245;  
Best Local Similarity 44.4%; Pred. No. 0.044;  
Matches 16; Conservative 11; Mismatches 8; Indels 1; Gaps 1;

QY 3 ESDTDPHFQDALMQLAKAVASAAALVTKAKSVAQ 36  
DB 38 EADESLNFEQIIEPAKSIATAATSLV-KAASAAQ 72

RESULT 15  
US-10-043-487-309  
Sequence 309, Application US/10043487  
Publication No. US20030055220A1  
GENERAL INFORMATION:  
APPLICANT: HYBRIGENICS  
APPLICANT: PIERRE, LEGRAIN  
TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptides  
FILE REFERENCE: B4778A  
CURRENT APPLICATION NUMBER: US/10/043,487  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/261,130  
PRIOR FILING DATE: 2001-01-12  
NUMBER OF SEQ ID NOS: 561  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 309  
LENGTH: 698  
TYPE: PRT  
ORGANISM: Shigella flexneri  
US-10-043-487-309

Query Match 40.6%; Score 73.5; DB 14; Length 698;  
Best Local Similarity 44.4%; Pred. No. 0.16;  
Matches 16; Conservative 11; Mismatches 8; Indels 1; Gaps 1;

QY 3 ESDTDPHFQDALMQLAKAVASAAALVTKAKSVAQ 38  
DB 491 EADESLNFEQIIEPAKSIATAATSLV-KAASAAQ 525

Search completed: November 10, 2004, 16:35:42  
Job time : 37.6512 secs

*This Page Blank (uspto)*



Db 149 PHLRRTKRIAKELASINALEPLNAN 175

### RESULT 3

T30177

Cytoskeleton assembly control protein homolog SLA2 - yeast (*Yarrowia lipolytica*)

C:Species: *Yarrowia lipolytica*, *Candida lipolytica*

C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T30177

R:Gausmann, U.; Schilabel, M.B.; Kurischko, C.

Submitted to the EMBL Data Library, July 1996

A:Description: SLA2 homologue of *Yarrowia lipolytica*.

A:Reference number: 220763

A:Accession: T30177

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1054 <GAP>

A:Cross-references: UNIPROT:Q94097, EMBL:U65409, NID:G3978133, PID:G3978134, PIDN:AA0831

A:Experimental source: strain W29

C:Genetics:

A:Gene: SLA2

A:Map position: III, adjacent to MATA

### Query Match

Best Local Similarity 30.1%; Score 54.5; DB 2; Length 1054;

Matches 15; Conservative 8; Mismatches 11; Indels 1; Gaps 1;

QY 3 ESPTDPHFDALMQLAKAVASAAALVLRKASVAQ 37

Db 841 KSTDLQHEAIAQAQAVINIAALI-RAATDAQ 874

### RESULT 4

T43195

Cytoskeleton assembly control protein SLA2 homolog - fission yeast (*Schizosaccharomyces*

C:Species: *Schizosaccharomyces pombe*

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000

C:Accession: T43195

R:Yoshioke, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.

DNA Res. 4, 363-369, 1997

A:Title: Identification of open reading frames in *Schizosaccharomyces pombe* cDNAs.

A:Reference number: 217323; PMID:98162722; PMID:9501991

A:Accession: T43195

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-468 <YOS>

A:Cross-references: EMBL:D89267, NID:G1749741, PIDN:BA13928.1; PID:G1749742

A:Experimental source: strain PR745

### Query Match

Best Local Similarity 29.8%; Score 54; DB 2; Length 468;

Matches 14; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 11 QDALMQLAKAVASAAALVLRKASVA 36

Db 386 ODLEDAKAVAEACALVQVESVA 411

### RESULT 5

G84218

Cysteine synthase [imported] - *Halobacterium* sp. NRC-1

C:Species: *Halobacterium* sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: G84218

R:Ng, W.V.; Kennedy, S.P.; Mahata, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

; Leichter, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jbdic

ung, K.H.; Alam, M.; Frettas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of *Halobacterium* species NRC-1.

A:Reference number: A84160; PMID:20504483; PMID:11016950

A:Accession: G84218

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-274 <STO>

A:Cross-references: UNIPROT:Q9HRP3, GB:AE004437, NID:G10580199, PIDN:AA01915.1; GSPDB:

C:Genetics:

A:Gene: YTHA

C:Superfamily: threonine dehydratase

### Query Match

Best Local Similarity 29.3%; Score 53; DB 2; Length 274;

Matches 15; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 3 ESPTDPHFDALMQLAKAVASAAALVLRKASVAQ 38

Db 201 DEDTHAEQRLAEQAQVASSAANSALADVAER 236

### RESULT 6

AD2842

Bacteriophytochrome protein [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)

C:Species: *Agrobacterium tumefaciens*

C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 25-Aug-2003

C:Accession: AD2842

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,

erage, G.; Gillet, M.; Grant, C.; Guenther, D.; Kutyavan, T.; Levy, R.; Li, M.; McChlel

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; PMID:21608550; PMID:11743193

A:Accession: AD2842

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-850 <KUR>

A:Cross-references: GB:AE008668; PIDN:AA143154.1; PID:G17740630; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atuz165

A:Map position: circular chromosome

C:Superfamily: bacteriophytochrome with receiver domain

### Query Match

Best Local Similarity 29.3%; Score 53; DB 2; Length 850;

Matches 14; Conservative 7; Mismatches 13; Indels 2; Gaps 1;

QY 2 GESDTPHFDALMQLAKAVASAAALVLRKASVAQ 37

Db 366 GHGATPPH-DALPRLARFVASBGRVATHTALSO 359

### RESULT 7

E97619

probable sensor/response regulator hybrid (PA1243) [imported] - *Agrobacterium tumefaciens*

C:Species: *Agrobacterium tumefaciens*

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 25-Aug-2003

C:Accession: E97619

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman

A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tu*

A:Reference number: A97359; PMID:21608551; PMID:11743194

A:Accession: E97619

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-850 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK87910.1; PID:G15157306; GSPDB:GN00169

C:Genetics:

A:Gene: AGR\_C\_3927

A:Map position: circular chromosome

C:Superfamily: bacteriophytochrome with receiver domain

### Query Match

Best Local Similarity 29.3%; Score 53; DB 2; Length 850;

Matches 14; Conservative 7; Mismatches 13; Indels 2; Gaps 1;



```

A:Accession: AH2669
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-571 <XLR>
A:Cross-references: UNIPROT:Q8UHC4, GB:AB008688, P1DN:AAL4774.1, P1D:g17739127, GSPDB
C:Genetics:
A:Experimental source: Strain C58 (Dupont)
A:Gene: dnaJ
A:Map position: circular chromosome

Query Match      29.0% Score 52.5; DB 2; Length 371;
Best Local Similarity 40.5% Pred. No. 22;
Matches 15; Conservative 4; Mismatches 17; Indels 1; Gaps 1;

OY      1 GGESDTPHFQDALMQLAKAVASAAAALVLRKASVAQ 37
          |||::|||::|||::|||::|||::|||::|||
Db      135 GADAQNDPKVQGA-ABAAKAASAKGDIPIAAGDTAQ 170

RESULT 13
C70906
hypochemical protein RV0184 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: C70906
R:COLE, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
R.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70906
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-249 <COL>
A:Cross-references: UNIPROT:O07428, GB:Z97050; GB:AL123456; NID:g3256008; P1DN:CAB09735
A:Experimental source: Strain H37RV
C:Genetics:
A:Gene: RV0184

Query Match      28.7% Score 52; DB 2; Length 249;
Best Local Similarity 38.2% Pred. No. 17;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY      3 ESDTDPHFQDALMQLAKAVASAAAALVLRKASVA 36
          |||::|||::|||::|||::|||::|||::|||
Db      18 EGTNDPHEADAFMSTAOQLATPAASIDLAVARSHA 51

RESULT 14
T35371
probable Na+/H+ antiporter - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T35371
R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, June 1999
A:Reference number: Z21576
A:Accession: T35371
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-514 <MR>
A:Cross-references: UNIPROT:Q9XAJ9; EMBL:AL079348; P1DN:CAB45470.1; GSPDB:GN00070; SCOREDB:SC66T3.14C
C:Genetics:
A:Gene: SCOREDB:SC66T3.14C

Query Match      28.7% Score 52; DB 2; Length 514;
Best Local Similarity 48.3% Pred. No. 37;
Matches 14; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

OY      2 GESDTPHFQDALMQLAKAVASAAAALVLR 30

```



Db 46 GEDGGLRFDPA--QLAQLGAALALIL 72

## RESULT 15

A44118  
annexin I type 2 - pigeon  
N:Alternate names: calpactin II; calpactin/lipocortin homolog; cp37 protein; crosnac 37k  
C:Species: Columba livia (domestic pigeon)  
C>Date: 31-Dec-1993 #sequence revision 26-May-1994 #text\_change 09-Jul-2004  
C:Accession: A44118; B44118; A38909; I5112  
R:Haigler, H.T.; Mangili, J.A.; Gao, Y.; Jones, J.; Horseman, N.D.  
U: Biol. Chem. 267, 19123-19129, 1992  
A>Title: Identification and characterization of columbid annexin Icp37. Insights into th  
A:Reference number: A44118; MUID:92406850; PMID:1388165  
A:Accession: A44118  
A:Molecule type: mRNA  
A:Residues: 1-343 <HA1>  
A:Cross-references: UNIPROT:Q92040  
A>Note: sequence extracted from NCBI backbone (NCBIN:113807, NCBIPI:113808)  
A:Accession: B44118  
A:Molecule type: protein  
A:Residues: 13-15, 'X', 17, 'X', 19-28, 'X', 30-50 <HA2>  
R:Haigler, H.T.; Mangili, J.A.; Gao, Y.; Jones, J.; Horseman, N.D.  
submitted to GenBank, October 1992  
A:Reference number: A38909  
A:Accession: A38909  
A:Molecule type: mRNA  
A:Residues: 1-73, 'HR', 76-243 <HA3>  
A:Cross-references: GB:M91008; NID:9213521; PIDN:AAA49447.1; PID:9213522  
A>Note: amino end is blocked in the mature form  
R:Gao, Y.; Horseman, N.D.  
Gene 143, 179-186, 1994  
A>Title: Structural and functional divergences of the columbid annexin I-encoding cp37 a  
A:Reference number: I5112; MUID:94265150; PMID:8206371  
A:Accession: I5112  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-154 <GAO>  
A:Cross-references: GB:L02504; NID:9213535; PIDN:AAA20674.1; PID:95239725  
A:Gene: cp37  
C:Genetics:  
A:Introns: 22/3; 55/1; 86/3; 124/3  
C:Superfamily: annexin I; annexin repeat homolog  
C:Keywords: acetylated amino end; calcium binding; duplication; endonexin fold; inflame  
F:2-342/Product: annexin I #status predicted <AX1>  
F:41-112/Domain: annexin repeat homolog <AX1>  
F:52-68/Region: endonexin fold #status predicted  
F:113-184/Domain: annexin repeat homolog <AX2>  
F:124-140/Region: endonexin fold #status predicted  
F:196-268/Domain: annexin repeat homolog <AX3>  
F:208-224/Region: endonexin fold #status predicted  
F:272-343/Domain: annexin repeat homolog <AX4>  
F:283-299/Region: endonexin fold #status predicted  
F:2/Modified site: acetylated amino end (Ala) (in mature form) #status predicted  
F:21/Binding site: phosphate (TYR) (covalent) (by EGF receptor/kinase) #status predicted  
F:24/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 28.2% Score 51; DB 1; Length 343;

Best Local Similarity 68.8%; Pred. No. 33;  
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 4 SDTDPHFQDALMQLAK 19  
Db 166 SPTSGHFQKALVYLAK 181

Search completed: November 10, 2004, 12:29:06  
Job time : 9.6821 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:27:01 ; Search time 41.5772 Seconds  
(without alignments)  
525.871 Million cell updates/sec

Title: US-10-092-750-6

Sequence: 1 GGBSPTDPHFDALMQLAKAVASAAALVLRKSVAGR 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trcml: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	175	96.7	2541	1	TN1_HUMAN
2	175	96.7	2541	1	O6Y490 homo sapien
3	175	96.7	2550	2	O86YD0
4	171	94.5	2564	2	O9UPX3
5	166	91.7	2541	1	O80TM2
6	156	86.2	2541	1	TN1_MOUSE
7	127	70.2	2350	2	O8AW10
8	127	70.2	2350	2	O71IX4
9	127	70.2	2542	1	AAO05019
10	87	48.1	2717	2	TN2_HUMAN
11	75	41.4	2816	2	O7OJ33
12	64.5	35.6	529	2	O9VSI8
13	64.5	35.6	529	2	O9PBB6
14	64.5	35.6	547	2	AAH59856
15	64.5	35.6	1452	2	O8CDM9
16	64.5	35.6	1471	2	O8CHG4
17	61	33.7	1527	2	O8MTA5
18	60	33.1	1584	2	O9VE55
19	59.5	32.9	1601	2	O860C2
20	58	32.0	996	2	O95XN3
21	58	32.0	2553	2	O22860
22	56	30.9	480	2	O747D6
23	56	30.9	480	2	AA36721
24	56	30.9	530	2	O87Y87
25	55	30.4	420	2	O9KYR1
26	55	30.4	423	1	YV2_CAEEL
27	54.5	30.1	1053	2	O6CCQ4
28	54.5	30.1	1054	2	O94097
29	54.5	30.1	1054	2	CAE84421
30	54	29.8	366	2	O7O9R8
31	54	29.8	389	2	O7KR74

32	54	29.8	389	2	AA222269	AA222269 drosophila
33	54	29.8	424	2	O86N74	O86N74 drosophila
34	54	29.8	471	2	O8PHV6	O8PHV6 xanthomonas
35	54	29.8	653	2	O8IXM8	O8IXM8 homo sapien
36	54	29.8	656	2	O7QRJ6	O7QRJ6 giardia lam
37	54	29.8	677	2	O6HP67	O6HP67 bacillus th
38	54	29.8	1073	2	O82EJ3	O82EJ3 streptomyc
39	54	29.8	1092	1	SLA2_SCHPO	O9P615 schistosach
40	53.5	29.6	199	2	O9E1X4	O9E1X4 cercopithe
41	53	29.3	274	2	O9HRP3	O9HRP3 halodacteri
42	53	29.3	378	2	O77CD9	O77CD9 bovine hepr
43	53	29.3	378	2	O65822	O65822 bovine hepr
44	53	29.3	850	2	O7CXQ4	O7CXQ4 agrobacteri
45	53	29.3	850	2	O8UDG1	O8UDG1 agrobacteri

## ALIGNMENTS

```

RESULT 1
ID TN1_HUMAN STANDARD; PRT; 2541 AA.
AC O9Y490; O9NZ02; Q9UHH8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Talin 1.
GN Name=TN1; Synonyms=TLN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RA SEQUENCE FROM N.A.
RA Mao L., Pan Y.H.;
RT "Complete cDNA sequence of human talin.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=20079169; PubMed=10610730;
RA Ben-Yosef T., Francomano C.A.;
RT "Characterization of the human talin (TN1) gene: genomic structure,
RL Genomics 62:316-319(1999).";
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal
CC structures to the plasma membrane. High molecular weight
CC cytoskeletal protein concentrated at regions of cell-substratum
CC contact and, in lymphocytes, at cell-cell contacts (By
CC similarity).
CC -!- SUBUNIT: Binds with high affinity to vinculin and with low
CC affinity to integrins (By similarity).
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -!- SIMILARITY: Contains 1 I/LWEQ domain.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
EMBL AF078828; AAD13152.1; -
EMBL AF177198; AAF23322.1; -
EMBL AF178534; AAF27330.1; -
EMBL AF178081; AAF27330.1; JOINED.
CC OGP; O9Y490; -
DR Genew; HGNC:11845; TN1.
DR MIM; 186745; -
CC GO; GO:0005925; C:focal adhesion; NAS.
DR GO; GO:0005900; P:structural constituent of cytoskeleton; NAS.
DR GO; GO:0005929; P:cell motility; NAS.
DR GO; GO:0007016; P:cytoskeletal anchoring; NAS.

```



```

Query Match          96.7%; Score 175; DB 2; Length 2550;
Best Local Similarity 100.0%; Pred. No. 1.9e-13;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GESDTPHFDALMOLAKAVASAAALVTKAKSVQR 38
Db 665 GESDTPHFDALMOLAKAVASAAALVTKAKSVQR 701

RESULT 4
O80TM2 PRELIMINARY; PRT; 2564 AA.
AC O80TM2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE MKIA1027 protein (Fragment).
GN Name=mkIA1027;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=22579291; PubMed=12693553;
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
DR EMBL; AK122420; BAC65702.4;
DR GO; GO:0005925; C:focal adhesion; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007044; P:cell-substrate junction assembly; IMP.
DR GO; GO:0030866; P:cortical actin cytoskeleton organization an. .; IMP.
FT NON TER
SO SEQUENCE 2564 AA; 272130 MW; B702BD310A5E283 CRC64;

Query Match          94.5%; Score 171; DB 2; Length 2564;
Best Local Similarity 97.3%; Pred. No. 6.4e-13;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GESDTPHFDALMOLAKAVASAAALVTKAKSVQR 38
Db 679 GESDTPHFDALMOLAKAVASAAALVTKAKSVQR 715

RESULT 5
TIN1 MOUSE STANDARD; PRT; 2541 AA.
AC P26039; O8VEP0;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Talin 1.
GN Name=Talin; Synonyms=Tln;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=91015390; PubMed=2120593;
RP REES D.J.G., Ades S.A., Singer S.J., Hynes R.O.;
RT "Sequence and domain structure of talin.";
RL Nature 347:685-689(1990).
FT (2)
RN RP SEQUENCE OF 1603-2541 FROM N.A.
RC TISSUE=Mammary gland;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Deje J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Mansina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stetson W., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein W.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Probably involved in connections of major cytoskeletal
CC structures to the plasma membrane. High molecular weight
CC cytoskeletal protein concentrated at regions of cell-substratum
CC contact and, in lymphocytes, at cell-cell contacts.
CC -1- SUBUNIT: Binds with high affinity to vinculin and with low
CC affinity to integrins.
CC -1- PTM: Phosphorylated.
CC -1- SIMILARITY: Contains 1 FERM domain.
CC -1- SIMILARITY: Contains 1 I/LWEQ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56123; CA39588.1;
DR EMBL; BC018557; AA018557.1;
DR PIR; S11661; S11661.
DR MGJ; MGJ:1099832; Tln.
DR GO; GO:0005925; C:focal adhesion; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007044; P:cell-substrate junction assembly; IMP.
DR GO; GO:0030866; P:cortical actin cytoskeleton organization an. .; IMP.
DR InterPro; IPR011000; Apopto_rtl_like.
DR InterPro; IPR010029; Band_4.1.
DR InterPro; IPR000299; Band_4.1.
DR InterPro; IPR009065; FERM.
DR InterPro; IPR002558; I/LWEQ.
DR InterPro; IPR011036; PH related.
DR Pfam; PF00373; Band_41.1.
DR Pfam; PF01608; I_LWEQ.1.
DR ProDom; PD011820; I/LWEQ.1.
DR ProSite; PS00660; FERM_1; 1.
DR ProSite; PS00661; FERM_2; 1.
DR ProSite; PS50057; FERM_3; 1.
DR ProSite; PS50945; I_LWEQ.1.
DR Cytoskeleton; Phosphorylation; Structural protein.
FT DOMAIN 86 403 FERM.
FT DOMAIN 2340 2533 I/LWEQ.
FT VARIANT 1105 1105 L -> P.
FT VARIANT 2180 2180 K -> W.
SO SEQUENCE 2541 AA; 269832 MW; 14EF75ABE9FC2CB CRC64;

Query Match          91.7%; Score 166; DB 1; Length 2541;
Best Local Similarity 94.6%; Pred. No. 2.8e-12;
Matches 35; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 GESDTPHFDALMOLAKAVASAAALVTKAKSVQR 38
Db 656 GESDTPHFDALMOLAKAVASAAALVTKAKSVQR 692

```

```

RESULT 6
Q8AW10 PRELIMINARY; PRT; 2541 AA.
ID Q8AW10;
AC Q8AW10;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Talin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97092324; PubMed=8937989;
RA Hemmings L., Rees D.J., Ochanian V., Bolton S.J., Gilmore A.P.,
RA Patel B., Fridde H., Trevithick J.E., Hynes R.O., Critchley D.R.;
RT "Talin contains three actin-binding sites each of which is adjacent to
RT a vinculin-binding site."
RN J Cell Sci. 109:2715-2726(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Hynes R.O.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY150847; AAT75275.1; -.
DR PIR; D42965; D42965.
DR PDB; 1MTX; X-ray; A=195-400.
DR PDB; 1MTZ; X-ray; B=200-400.
DR PDB; 1MK7; X-ray; B/D=209-400.
DR PDB; 1MK9; X-ray; B/D/F/H=209-400.
DR GO; GO:0005856; Cytoskeleton; IEA.
DR GO; GO:0003779; F-actin binding; IEA.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR000299; Band_4.1.
DR InterPro; IPR009065; FERM.
DR InterPro; IPR002558; ILMEQ.
DR InterPro; IPR011036; PH-related.
DR Pfam; PF00373; Band_4.1.
DR Pfam; PF01608; I_ILMEQ; 1.
DR Prodom; PD011820; ILMEQ; 3.
DR SMART; SM00295; B41; 1.
DR SMART; SM00307; ILMEQ; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS0057; FERM_3; 1.
DR PROSITE; PS00945; I_ILMEQ; 1.
SQ SEQUENCE 2541 AA; 271839 MW; 5A94C290C624699E CRC64;

Query Match 86.2%; Score 156; DB 2; Length 2541;
Best Local Similarity 89.2%; Pred. No. 5,6e-11;
Matches 33; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GESDTPHFODALMOLAKAVASAAALVYKAKSVQ 38
DB 656 GESDTPHFODALMOLAKAVASAAALVYKAKSVQ 692

RESULT 7
Q7ILX4 PRELIMINARY; PRT; 2350 AA.
ID Q7ILX4;
AC Q7ILX4;
DT 05-JUL-2004 (TREMblrel. 27, Created)
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)
DE Talin 2 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Dubois A., Albiges-Rizo C., Block M., Faessler R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467081; AAQ05019.1; -.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR000299; Band_4.1.
DR InterPro; IPR009065; FERM.
DR InterPro; IPR002558; ILMEQ.
DR InterPro; IPR011036; PH-related.
DR Pfam; PF00373; Band_4.1.
DR Pfam; PF01608; I_ILMEQ; 1.
DR Prodom; PD011820; ILMEQ; 2.
DR SMART; SM00295; B41; 1.
DR SMART; SM00307; ILMEQ; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS0057; FERM_3; 1.
DR PROSITE; PS00945; I_ILMEQ; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 2350 AA; 250820 MW; 0188E1663846CF14 CRC64;

Query Match 70.2%; Score 127; DB 2; Length 2350;
Best Local Similarity 75.0%; Pred. No. 3e-07;
Matches 27; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 GESDTPHFODALMOLAKAVASAAALVYKAKSVQ 37
DB 635 GESDTPHFODALMOLAKAVASAAALVYKAKSVQ 670

RESULT 8
AAQ05019 PRELIMINARY; PRT; 2350 AA.
ID AAQ05019;
AC AAQ05019;
DT 02-MAR-2004 (TREMblrel. 27, Created)
DT 02-MAR-2004 (TREMblrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMblrel. 27, Last annotation update)
DE Talin 2 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Dubois A., Albiges-Rizo C., Block M., Faessler R.;
RL "Expression of the newly identified Mus musculus talin 2 gene."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467081; AAQ05019.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 2350 AA; 250820 MW; 0188E1663846CF14 CRC64;

Query Match 70.2%; Score 127; DB 2; Length 2350;
Best Local Similarity 75.0%; Pred. No. 3e-07;
Matches 27; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 GESDTPHFODALMOLAKAVASAAALVYKAKSVQ 37
DB 635 GESDTPHFODALMOLAKAVASAAALVYKAKSVQ 670

RESULT 9
TIN2_HUMAN STANDARD; PRT; 2542 AA.
ID TIN2_HUMAN;
AC Q9Y466;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Talin 2.
GN Name=TIN2; Synonyms=KIAA0320;
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Skeletal muscle;  
 RA MEDLINE=22309504; PubMed=14422219; DOI=10.1038/nature01147;  
 DI Paolo G., Pellegrini L., Letinic K., Cestra G., Zoncu R.,  
 Voronov S., Chang S., Guo J., Wenk M.R., De Camilli P.,  
 "Recruitment and regulation of phosphatidylinositol phosphate kinase  
 RT type I gamma by the FERM domain of talin.";  
 RL Nature 420:85-89(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21418781; PubMed=11527381; DOI=10.1006/birc.2001.5497;  
 RA Monkley S.J., Pritchard C.A., Critchley D.R.,  
 "Analysis of the mammalian talin2 gene TM2.";  
 RL Biochem. Biophys. Res. Commun. 286:880-885(2001).  
 RN [3]  
 RP SEQUENCE OF 610-2542 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=97349984; PubMed=9205841;  
 RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,  
 Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.,  
 "Prediction of the coding sequences of unidentified human genes. VII.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RL code for large proteins in vitro.";  
 CC DNA Res. 4:141-150(1997).  
 CC -1- FUNCTION: As a major component of focal adhesion plaques that  
 CC links integrin to the actin cytoskeleton, may play an important  
 CC role in cell adhesion. Recruits PI5K1C to focal adhesion plaques  
 CC and strongly activates its kinase activity (By similarity).  
 CC -1- SUBUNIT: Interacts directly with PI5K1C.  
 CC -1- SUBCELLULAR LOCATION: Focal adhesion plaques and synapses.  
 CC -1- SIMILARITY: Contains 1 FERM domain.  
 CC -1- SIMILARITY: Contains 1 ILMEQ domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation on  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AF402000; AAM73764.1; -;  
 DR EMBL; AB002318; BAA20778.2; -;  
 DR GenBank; HGNC:15447; TM2.  
 DR MIM; 607348; -;  
 DR GO; GO:0015629; C:actin cytoskeleton; NAS.  
 DR GO; GO:0045202; C:synapse; NAS.  
 DR GO; GO:0003779; F:actin binding; NAS.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR GO; GO:0005198; F:structural molecule activity; NAS.  
 DR GO; GO:0007155; P:cell adhesion; NAS.  
 DR InterPro; IPR011000; Apolp\_TITL\_like.  
 DR InterPro; IPR000299; Band\_4.1.1.  
 DR InterPro; IPR009065; FERM.  
 DR InterPro; IPR002558; ILMEQ.  
 DR InterPro; IPR01036; PH related.  
 DR Pfam; PF00373; Band\_4.1; 1.  
 DR Pfam; PF01608; I\_ILMEQ; 1.  
 DR ProDom; PD011820; ILMEQ; 1.  
 DR ProSITE; PS00660; FERM\_1; FALSE\_NEG.  
 DR ProSITE; PS00661; FERM\_2; 1.  
 DR ProSITE; PS00577; FERM\_3; 1.  
 DR ProSITE; PS00945; I\_ILMEQ; 1.  
 KM Cytoskeleton; Polymorphism; Structural protein.  
 FT DOMAIN 68 406  
 FT DOMAIN 2341 2533  
 FT DOMAIN 312 406  
 FT DOMAIN 861 1193  
 FT VARIANT 1148 1148 A -> S (in dbSNP:2280279).

FT VARIANT 2266 2266 /FTID=VAR\_014432.  
 FT CONFLICT 1269 1269 F -> L (in dbSNP:3816988).  
 FT CONFLICT 2542 AA; 271552 MM; 0A83FD26D914D41 CRC64; /FTID=VAR\_014433.  
 SQ SEQUENCE 2542 AA; 271552 MM; 0A83FD26D914D41 CRC64; L -> F (in Ref. 1).  
 Query Match 70.2%; Score 127; DB 1; Length 2542;  
 Best Local Similarity 75.0%; Pred. No. 3.2e-07;  
 Matches 27; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 GESDTPHFODALMOLAKAVASAAALVLRKXSVQ 37  
 DB 659 GENETDERFODVLMSLAKAVANAAALVLRKXSVQ 694  
 RESULT 10  
 ID Q7QJ3 PRELIMINARY; PRT; 2717 AA.  
 AC Q7QJ3;  
 DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE AGP3777.  
 GN Name=egCG56699; ORFNames=ENSG0000001618;  
 OS Anopheles gambiae str. PBST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PBST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC -----  
 DR EMBL; AAB0108807; EAA04618.1; -;  
 DR GO; GO:0005856; C:cytoskeleton; IEA.  
 DR GO; GO:0003779; F:actin binding; IEA.  
 DR InterPro; IPR000299; Band\_4.1.  
 DR InterPro; IPR002558; ILMEQ.  
 DR Pfam; PF00373; Band\_4.1; 1.  
 DR Pfam; PF01608; I\_ILMEQ; 1.  
 DR ProDom; PD011820; ILMEQ; 1.  
 DR ProSITE; PS00660; FERM\_1; UNKNOWN\_1.  
 DR ProSITE; PS00577; FERM\_3; 1.  
 DR ProSITE; PS00945; I\_ILMEQ; 1.  
 DR PROSITE; PS00945; I\_ILMEQ; 1.  
 SQ SEQUENCE 2717 AA; 292666 MM; 360F17B59AB83F4D CRC64;  
 Query Match 48.1%; Score 87; DB 2; Length 2717;  
 Best Local Similarity 58.3%; Pred. No. 0.052;  
 Matches 21; Conservative 3; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 GGSDDTPHFODALMOLAKAVASAAALVLRKXSVQ 36  
 DB 674 GGSVSRRLHMLGAKAVANTTALVLRKXSVQ 709  
 RESULT 11  
 ID Q9VSL8 PRELIMINARY; PRT; 2836 AA.  
 AC Q9VSL8;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE CG6831-PA (Talin).  
 GN Name=Talin; ORFNames=CG6831;  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams W.D., Celniker S.E., Holt R.A., Evans C.A., Goeysen J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazef R.G., Champagne M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.Y., Andrews-Frankoch C., Baldwin D.,
RA Bailew R.M., Basu P.V., Bendale U., Bayraktaroglu L., Beasly E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu B., Center A., Chandra I.,
RA Chery J.M., Cawley S., Dalike C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes W., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Klup D., Lai Z.,
RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett M., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng X.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams W., Champagne M., Dugan S.P., Frise E., Hogson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svitek R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svitek R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.,
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomic perspective";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,
RA Betancourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.,
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review";

```

```

RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Brody T., Kuzin A., Stivers C., Odenwald M.F.,
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003555; AAF5039.1; -
DR EMBL; AF29248; AAG22814.1; -
DR FLYBASE; FBgn0020254; rhea.
DR GO; GO:0005856; Cytoskeleton; IEA.
DR GO; GO:0003779; F-actin binding; IEA.
DR InterPro; IPR000299; Band_4.1.
DR InterPro; IPR009065; FERM.
DR InterPro; IPR002558; ILMEO.
DR InterPro; IPR011036; PH related.
DR Pfam; PF00373; Band_4; 1.
DR Pfam; PF01608; ILMEO; 1.
DR SMART; SM00295; B4; 1.
DR PROSITE; PS00660; FERM_1; UNKNOWN_1.
DR PROSITE; PS50057; FERM_3; 1.
DR PROSITE; PS50945; ILMEO; 1.
SQ SEQUENCE 2836 AA; 306391 MW; 11A52A01CD3683B8 CRC64;

Query Match 41.4%; Score 75; DB 2; Length 2836;
Best Local Similarity 64.0%; Pred. No. 2;
Matches 16; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 12 DALMOLAKAVASNAALVAKASVA 36
Db 684 DMLALAKAVANTTALVAKXIA 708

RESULT 12
ID 06PB69 PRELIMINARY; PRT; 529 AA.
AC 06PB69;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Tln2 protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenner C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins K.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Diatchenko L., Matovina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Caesvant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchan J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;

```



```

RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences."
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6; TISSUE=Brain;
RA      Strausberg R.;
RL      Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC059836; AAH59856.1; .
DR      InterPro; IPR011000; Apolp_III_1like.
DR      Pfam; PF01608; I_LMEQ; 1.
DR      ProDom; PD01820; ILMEQ; 1.
DR      SMART; SM00307; ILMEQ; 1.
DR      PROSITE; PS50945; I_LMEQ; 1.
FT      NON_TER
SQ      SEQUENCE 529 AA; 56024 MW; 5E84BA89975D1CBC CRC64;

Query Match      35.6%; Score 64.5; DB 2; Length 529;
Best Local Similarity 38.9%; Pred. No. 9.2;
Matches 14; Conservative 12; Mismatches 9; Indels 1; Gaps 1;

QY      3 ESDTDFHODALMOLAKAVASAAALVTKKSVQR 38
DB      322 QADETLDFEEOITLFAKSIATATSLV-KSASAAQR 356

RESULT 13
AAH59856 PRELIMINARY; PRT; 529 AA.
ID AAH59856;
AC AAH59856;
DT 02-MAR-2004 (TRENBLrel. 27, Created)
DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)
DE Tln2 protein (Fragment).
DE Tln2.
GN Tln2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=1477932;
RA Strausberg R.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner K.H.; Schaefer C.F.; Bhat N.K.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Sherman C.M.; Schuler G.D.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA Ditchenko L.; Marnett K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stepleton M.; Soares W.B.; Bonaldi M.F.; Casavant T.L.; Scheetz T.E.;
RA Brownstein W.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA Rata S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.;
RA Bosak S.A.; McEwen P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulik S.W.;
RA Villalon D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA Fanev J.; Helton E.; Kettman M.; Madan A.; Rodriguez S.; Sanchez A.;
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Butcherfield Y.S.;
RA Krzyzanski M.I.; Skalski U.; Smalusz D.E.; Scherch A.; Schein J.E.;
RA Jones S.J.; Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC059856; AAH59856.1; .
FT NON_TER
SQ SEQUENCE 529 AA; 56024 MW; 5E84BA89975D1CBC CRC64;

```

```

Query Match      35.6%; Score 64.5; DB 2; Length 529;
Best Local Similarity 38.9%; Pred. No. 9.2;
Matches 14; Conservative 12; Mismatches 9; Indels 1; Gaps 1;

QY      3 ESDTDFHODALMOLAKAVASAAALVTKKSVQR 38
DB      322 QADETLDFEEOITLFAKSIATATSLV-KSASAAQR 356

RESULT 14
ID Q9CS79 PRELIMINARY; PRT; 547 AA.
AC Q9CS79;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched
DE library, clone:5730421P04 product:caln 2, full insert sequence.
DE (Fragment).
GN Name=Tln2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Whole body;
RX MEDLINE=9279253; PubMed=10349536;
RA Carninci P.; Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Whole body;
RX MEDLINE=2049374; PubMed=11042159;
RA Carninci P.; Shibata Y.; Hayatsu N.; Sugahara Y.; Shibata K.; Itoh M.;
RA Kono H.; Okazaki Y.; Muramatsu M.; Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K.; Itoh M.; Aizawa K.; Nagaoka S.; Sasaki N.; Carninci P.;
RA Kono H.; Akiyama U.; Nishi K.; Kitsuai T.; Tashiro H.; Itoh M.;
RA Sumi N.; Ishii Y.; Nakamura S.; Hazama M.; Nishino T.; Harada A.;
RA Yamamoto R.; Matsumoto H.; Sakaguchi S.; Ikegami T.; Kashiwagi K.;
RA Fujiwaka S.; Inoue K.; Togawa Y.; Izawa M.; Ohara E.; Watahiki K.;
RA Yoneda Y.; Ishikawa T.; Ozawa K.; Tanaka T.; Matsura S.; Kawai J.;
RA Okazaki Y.; Muramatsu M.; Inoue Y.; Kira A.; Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Whole body;
RA Adachi J.; Aizawa K.; Akahira S.; Akimura T.; Arai A.; Aono H.;

```

```

RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Yamaguchi M., Tagawa Y., Tanaka T.,
RA Teijima Y., Toya T., Yamaguchi M., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK017597; BAB30830.1; -.
DR MGD: MGI:1917799; Tln2.
DR GO: GO:0015629; C:actin cytoskeleton; ISS.
DR GO: GO:0043202; C:synapse; ISS.
DR GO: GO:0003779; F:actin binding; ISS.
DR GO: GO:0005515; F:protein binding; ISS.
DR GO: GO:0005198; F:structural molecule activity; ISS.
DR GO: GO:0007155; P:cell adhesion; ISS.
DR InterPro: IPR011000; Apolp_III_like.
DR InterPro: IPR002558; LMEQ.
DR Pfam: PF01608; LMEQ; 1.
DR ProDom: PD011820; LMEQ; 1.
DR PROSITE: PS50945; LMEQ; 1.
FT NON TER
SQ SEQUENCE 547 AA; 58002 MW; F51702D23446EACA CRC64;

Query Match 35.6%; Score 64.5; DB 2; Length 547;
Best Local Similarity 38.9%; Pred. No. 9.5;
Matches 14; Conservative 12; Mismatches 9; Indels 1; Gaps 1;

OY 3 ESDTPHFQDALMQLAKAVASAAALVLRKSAVQR 38
Db 325 QADETLDFEQILBAKSIATATSLV-KSASAAQR 359

RESULT 15
Q8CDM9 PRELIMINARY; PRT; 1452 AA.
AC Q8CDM9;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE library, clone:4931402J01 product:italin 2, full insert sequence.
GN Name=Tln2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujisake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384 format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Aichi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohgato N., Okazaki Y.,
RA Saito K., Saitoh K., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahata S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK029828; BAC26635.1; -.
DR MGD: MGI:1917799; Tln2.
DR GO: GO:0015629; C:actin cytoskeleton; ISS.
DR GO: GO:0043202; C:synapse; ISS.
DR GO: GO:0003779; F:actin binding; ISS.
DR GO: GO:0005515; F:protein binding; ISS.
DR GO: GO:0005198; F:structural molecule activity; ISS.
DR GO: GO:0007155; P:cell adhesion; ISS.
DR InterPro: IPR011000; Apolp_III_like.
DR InterPro: IPR002558; LMEQ.
DR Pfam: PF01608; LMEQ; 1.
DR ProDom: PD011820; LMEQ; 2.
DR PROSITE: PS50945; LMEQ; 1.
SQ SEQUENCE 1452 AA; 153347 MW; 0F626C0D650241E CRC64;

Query Match 35.6%; Score 64.5; DB 2; Length 1452;
Best Local Similarity 38.9%; Pred. No. 24;
Matches 14; Conservative 12; Mismatches 9; Indels 1; Gaps 1;

OY 3 ESDTPHFQDALMQLAKAVASAAALVLRKSAVQR 38
Db 1245 QADETLDFEQILBAKSIATATSLV-KSASAAQR 1279

Search completed: November 10, 2004, 12:26:52
Job time : 45.5772 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17, Search time 10.4244 Seconds  
(without alignments)  
222.664 Million cell updates/sec

Title: US-10-092-750-7  
Perfect score: 174  
Sequence: 1 GTRQDRMPTMAIEIEQLARLVGVNDKAEYTN 35

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 47813 segs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	32.2	343	4	US-09-543-681A-6459
2	55	31.6	81	4	US-09-621-976-7145
3	54.5	31.3	1463	4	US-09-583-166-4076
4	52.5	30.2	657	3	US-09-355-166-4076
5	52.5	30.2	657	4	US-09-462-845-2
6	50	28.7	387	4	US-09-107-532A-5175
7	50	28.4	1173	4	US-09-248-796A-19313
8	49.5	28.4	662	4	US-09-134-000C-5682
9	49	28.2	205	4	US-09-248-796A-1984
10	49	28.2	237	4	US-09-248-796A-1984
11	49	28.2	3210	4	US-09-538-092-1154
12	49	28.2	3248	5	US-08-353-700-1
13	49	28.2	3248	5	US-08-353-700-1
14	48.5	27.9	280	4	US-09-293-858B-8
15	48.5	27.9	301	4	US-09-270-767-3369
16	48.5	27.9	301	4	US-09-270-767-3369
17	48.5	27.9	812	4	US-09-270-767-3369
18	48.5	27.9	812	4	US-09-270-767-3369
19	48.5	27.9	1671	4	US-09-270-767-3369
20	48.5	27.9	1671	4	US-09-270-767-3369
21	47	27.0	1025	4	US-09-489-039A-8093
22	47	27.0	5024	4	US-09-710-279-2964
23	46.5	26.7	298	4	US-09-253-991A-2545
24	46	26.4	271	3	US-08-433-263B-2
25	46	26.4	286	4	US-09-248-796A-17854
26	46	26.4	335	1	US-08-348-891A-4
27	46	26.4	335	2	US-08-905-817-4

28	46	26.4	457	4	US-10-124-601-2	Sequence 2, Appli
29	46	26.4	526	3	US-08-504-878A-2	Sequence 2, Appli
30	46	26.4	526	4	US-09-466-517A-2	Sequence 2, Appli
31	46	26.4	549	4	US-09-693-146-2	Sequence 2, Appli
32	46	26.4	566	4	US-09-134-000C-4385	Sequence 4385, Ap
33	46	26.4	633	4	US-09-583-110-3803	Sequence 3803, Ap
34	45.5	26.1	198	4	US-09-248-796A-18306	Sequence 18306, A
35	45.5	26.1	2802	3	US-09-542-331-1	Sequence 1, Appli
36	45.5	26.1	2802	3	US-09-510-791-1	Sequence 1, Appli
37	45	25.9	93	4	US-09-513-999C-7703	Sequence 7703, Ap
38	45	25.9	209	4	US-09-599-360B-121	Sequence 121, App
39	45	25.9	225	4	US-09-107-532A-7302	Sequence 7302, Ap
40	45	25.9	231	4	US-09-583-110-3257	Sequence 3257, Ap
41	45	25.9	264	4	US-09-252-991A-18073	Sequence 18073, A
42	45	25.9	317	4	US-09-328-352-6391	Sequence 6391, Ap
43	45	25.9	338	4	US-09-489-039A-11745	Sequence 11745, A
44	45	25.9	446	3	US-09-457-040B-28	Sequence 28, Appli
45	45	25.9	472	4	US-09-540-236-3760	Sequence 3760, Ap

## ALIGNMENTS

RESULT 1  
US-09-543-681A-6459  
Sequence 6459, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543.681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 6459  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-6459

Query Match 32.2% Score 56; DB 4; Length 343;  
Best Local Similarity 48.1% Pred. No. 1.4;  
Matches 13; Conservative 7; Mismatches 5; Indels 2; Gaps 1;

QY 9 ETMAIEBOL-LARLVGVNDKAEYTN 33  
DB 137 DSTAIEICIDBIKLAGVCDKVAEIT 163  
RESULT 2  
US-09-621-976-7145  
Sequence 7145, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
TITLE OF INVENTION: ESTE and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 7145  
LENGTH: 81  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-621-976-7145  
Query Match 31.6% Score 55; DB 4; Length 81;

Best Local Similarity 45.8%; Pred. No. 0.35;  
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

CY 1 GTRODRMFMATIEBOLARLNG 24  
DB 58 GSSODRMEFTRRLNPNFQGLQ 81

RESULT 3  
US-09-583-110-4076

Sequence 4076, Application US/09583110  
Patent No. 6699703

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al.

TITLE OF INVENTION: Nucleic Acid Sequences Relating to Streptococcus

FILE REFERENCE: PAT400-07A

CURRENT FILING DATE: 2000-05-26

PRIOR FILING DATE: 1998-06-30

PRIOR FILING DATE: 1998-05-12

PRIOR FILING DATE: 1997-07-02

NUMBER OF SEQ ID NOS: 5322

SEQ ID NO 4076

LENGTH: 1463

TYPE: PRT

ORGANISM: Streptococcus pneumoniae

US-09-583-110-4076

Query Match 31.3%; Score 54.5; DB 4; Length 1463;  
Best Local Similarity 36.8%; Pred. No. 14;

Matches 14; Conservative 11; Mismatches 8; Indels 5; Gaps 3;

CY 2 TRODRMFMATIEBOLARLNG--VNDXMAEYNA 35

DB 237 TEENRIVEGVDFVEQKXTR-TGRVLFNFKMTDITSS 273

RESULT 4  
US-09-355-166-1

Sequence 1, Application US/09355166

Patent No. 6316241

GENERAL INFORMATION:

APPLICANT: Genencor International, Inc.

TITLE OF INVENTION: Alpha/Beta Hydrolyase-Fold Enzymes

FILE REFERENCE: GC511-PCT

CURRENT FILING DATE: 1999-07-20

NUMBER OF SEQ ID NOS: 21

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 657

TYPE: PRT

ORGANISM: Bacillus

US-09-355-166-1

Query Match 30.2%; Score 52.5; DB 3; Length 657;  
Best Local Similarity 42.4%; Pred. No. 11;

Matches 14; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

CY 2 TRODRMFMATIEBOLARLNGVNDK-MAEYT 33

DB 365 TKDRESELYSIPLGQEKQLTGANDKPREHT 397

RESULT 5  
US-09-462-845-2

Sequence 2, Application US/09462845

Patent No. 6723530

GENERAL INFORMATION:

APPLICANT: Estell, David A.

TITLE OF INVENTION: Proteases From Gram-Positive Organisms

FILE REFERENCE: GC382-US

CURRENT FILING DATE: 2000-01-13

PRIOR FILING DATE: 1998-07-14

PRIOR FILING DATE: 1998-07-15

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 657

TYPE: PRT

ORGANISM: Bacillus subtilis

US-09-462-845-2

Query Match 30.2%; Score 52.5; DB 4; Length 657;  
Best Local Similarity 42.4%; Pred. No. 11;

Matches 14; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

CY 2 TRODRMFMATIEBOLARLNGVNDK-MAEYT 33

DB 365 TKDRESELYSIPLGQEKQLTGANDKPREHT 397

RESULT 6  
US-09-107-532A-5175

Sequence 5175, Application US/09107532A

Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: Only 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5175:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (B) LOCATION 1..387  
SEQUENCE DESCRIPTION: SEQ ID NO: 5175;  
US-09-107-532A-5175

Query Match 28.7%; Score 50; DB 4; Length 387;  
Best Local Similarity 42.3%; Pred. No. 14;  
Matches 11; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 4 QDRMFEYMAI--EIQRLARLGVND 27  
DB 343 OSRQLEMKIHENHLSQLEVTGIND 368

RESULT 7  
US-09-248-796A-19313  
; Sequence 19313, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 19313  
; LENGTH: 1173  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (210)  
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknd  
US-09-248-796A-19313

Query Match 28.7%; Score 50; DB 4; Length 1173;  
Best Local Similarity 36.8%; Pred. No. 54;  
Matches 7; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 16 EQLARLGVNDKMAEYTN 34  
DB 906 DQLSKLNGINQELTKYSH 924

RESULT 8  
US-09-134-000C-5682  
; Sequence 5682, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5682  
; LENGTH: 662  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-5682

Query Match 28.4%; Score 49.5; DB 4; Length 662;  
Best Local Similarity 37.8%; Pred. No. 32;  
Matches 14; Conservative 6; Mismatches 12; Indels 5; Gaps 1;

QY 2 TRQDRMFEYMAIEIQRLARLT----GVNDKMAEYTN 33  
DB 62 TEQLRKYEYMLQQLQELHLLFLTRSSGKGLKIKRKRYTN 98

RESULT 9  
US-09-248-796A-19844  
; Sequence 19844, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 19844  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (180)  
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkr  
US-09-248-796A-19844

Query Match 28.2%; Score 49; DB 4; Length 205;  
Best Local Similarity 26.6%; Pred. No. 9.2;  
Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 5 DRMFETMAIEIQRLARLGVNDKMAEY 32  
DB 173 DRLLLELXLPKCEYKYGINDINDY 200

RESULT 10  
US-09-248-796A-19815  
; Sequence 19815, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 19815  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-19815

Query Match 28.2%; Score 49; DB 4; Length 237;  
Best Local Similarity 45.8%; Pred. No. 11;  
Matches 11; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 13 IEIQRLARL--TGVNDKMAEYTN 34  
DB 24 ISMRQLILQILITGTINKYKRYTN 47

RESULT 11  
US-09-538-092-1154

Sequence 1154, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Giot, Loic  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CurataseqFormatter Version 0.9  
SEQ ID NO 1154  
LENGTH: 3210  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: Polypeptide Accession Number P49454  
US-09-538-092-1154

Query Match 28.2%; Score 49; DB 4; Length 3210;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 TRODMFETMAIEBQLARLT 23  
DB 401 SRQGRSFQTLDECIQMKARLT 422

RESULT 12  
US-08-353-700-1  
Sequence 1, Application US/08353700  
Patent No. 5539919  
GENERAL INFORMATION:  
APPLICANT: YEN, TIMOTHY J.  
TITLE OF INVENTION: NUCLEIC ACID ENCODING A  
TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,  
TITLE OF INVENTION: AND METHODS OF USE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DANN, DOREMAN, HERRELL AND SKILLMAN  
STREET: 1601 MARKET STREET, SUITE 720  
CITY: PHILADELPHIA  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/353,700  
FILING DATE: 09-DEC-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: REED, JANET E.  
REGISTRATION NUMBER: 36,252  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 563-4100  
TELEFAX: (215) 563-4044  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3248 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: Protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: HUMAN  
US-08-353-700-1

Query Match 28.2%; Score 49; DB 1; Length 3248;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 TRODMFETMAIEBQLARLT 23  
DB 401 SRQGRSFQTLDECIQMKARLT 422

RESULT 13  
PCT-US95-16216-1  
Sequence 1, Application PC/TUS9516216  
GENERAL INFORMATION:  
APPLICANT: Yen, Timothy J.  
TITLE OF INVENTION: Nucleic Acid Encoding a Transiently  
TITLE OF INVENTION: Expressed Kinetochore Protein, and Methods of Use  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dann, Doreman, Herrell and Skillman  
STREET: 1601 Market Street Suite 720  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16216  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/353,700  
FILING DATE: 09-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Reed, Janet E.  
REGISTRATION NUMBER: 36,252  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 563-4100  
TELEFAX: (215) 563-4044  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3248 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PCT-US95-16216-1

Query Match 28.2%; Score 49; DB 5; Length 3248;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 TRODMFETMAIEBQLARLT 23  
DB 401 SRQGRSFQTLDECIQMKARLT 422

RESULT 14  
US-08-292-858B-8  
Sequence 8, Application US/09292858B

```

; Patent No. 6455681
; GENERAL INFORMATION:
; APPLICANT: Dean, Frank
; APPLICANT: O'Donnell, Michael E.
; TITLE OF INVENTION: DNA MOLECULES ENCODING SINGLE STRAND GAP RESPONSE
; TITLE OF INVENTION: PROTEINS INVOLVED IN ACTIVATION OF A DNA REPAIR/CELL
; TITLE OF INVENTION: CYCLE CHECKPOINT PATHWAY
; FILE REFERENCE: 22221/1011
; CURRENT APPLICATION NUMBER: US/09/292,858B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,020
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 8
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-292-858B-8

```

```

Query Match          27.9%; Score 48.5; DB 4; Length 280;
Best Local Similarity 34.8%; Pred. No. 16;
Matches 16; Conservative 5; Mismatches 10; Indels 15; Gaps 2;

```

```

QY      2 TRQDRMFETWA--IEIEQLRLRTG-----VNDKMAEY 32
DB      218 THEDRVVEMAEVHIDIRKLTQFLAGQGVNPTKALCNIVNNQVHF 263

```

```

RESULT 15
US-09-270-767-33698
; Sequence 33698, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 33698
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-33698

```

```

Query Match          27.9%; Score 48.5; DB 4; Length 301;
Best Local Similarity 35.5%; Pred. No. 18;
Matches 11; Conservative 8; Mismatches 9; Indels 3; Gaps 1;

```

```

QY      4 QDRMFETVAIEIEQLRLRTGVDKMAEYTN 34
DB      210 QDQIDKHAIVKIVKQRLQGITD--GDYDN 237

```

Search completed: November 10, 2004, 12:32:22  
Job time : 11.4244 secs

***This Page Blank (usp10)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 ; Search time 33.7577 Seconds  
(without alignments)  
366.225 Million cell updates/sec

Title: US-10-092-750-7  
Perfect score: 174  
Sequence: 1 GTRDPRMFTWAIETIQLRLTGVDKMAEYNTA 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35322586 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	174	100.0	35	US-10-092-750-7
2	77	44.3	57	US-09-925-299-1540
3	77	44.3	57	US-09-925-299-1540
4	57	32.8	241	US-10-424-559-148251
5	57	32.8	601	US-10-369-493-9711
6	54.5	31.3	1460	US-09-815-242-13668
7	54.5	31.3	1460	US-10-282-122A-73765
8	54	31.0	387	US-09-738-626-4288
9	54	31.0	1034	US-09-815-242-10331
10	52.5	30.2	657	US-09-950-366-1
11	52.5	30.2	657	US-10-401-437-2
12	52.5	30.2	657	US-10-402-312-2
13	52.5	30.2	657	US-10-402-067-2

14	52.5	30.2	657	14	US-10-401-436-2	Sequence 2, Appl1
15	52	29.9	345	16	US-10-437-963-192198	Sequence 192198,
16	51.5	29.6	1465	15	US-10-282-122A-74767	Sequence 74767, A
17	51	29.3	243	17	US-10-425-115-320335	Sequence 320335,
18	51	29.3	254	15	US-10-425-114-49112	Sequence 49112, A
19	51	29.3	553	14	US-10-104-047-2655	Sequence 2655, Ap
20	51	29.3	988	15	US-10-424-559-144895	Sequence 144895,
21	51	29.3	1044	15	US-10-425-114-72709	Sequence 72709, A
22	50	28.7	363	15	US-10-282-122A-57571	Sequence 57571, A
23	50	28.7	393	16	US-10-437-963-137447	Sequence 137447, A
24	50	28.7	703	17	US-10-739-930-10003	Sequence 10003, A
25	50	28.7	1240	14	US-10-032-585-7366	Sequence 7366, Ap
26	50	28.7	1616	16	US-10-437-963-166909	Sequence 166909,
27	50	28.7	1620	16	US-10-437-963-148708	Sequence 148708,
28	50	28.7	1670	16	US-10-437-963-148711	Sequence 148711,
29	50	28.7	1736	16	US-10-437-963-168748	Sequence 168748,
30	50	28.7	1780	16	US-10-437-963-148705	Sequence 148705,
31	49.5	28.4	666	15	US-10-424-559-231698	Sequence 231698,
32	49	28.2	230	16	US-10-451-467A-508	Sequence 508, App
33	49	28.2	247	16	US-10-437-963-146738	Sequence 146738,
34	49	28.2	333	15	US-10-425-114-52428	Sequence 52428, A
35	49	28.2	626	15	US-10-282-122A-71772	Sequence 71772, A
36	49	28.2	869	15	US-10-424-559-244705	Sequence 244705,
37	49	28.2	1510	15	US-10-282-122A-55505	Sequence 55505, A
38	49	28.2	1665	16	US-10-437-963-151216	Sequence 151216,
39	49	28.2	1706	16	US-10-437-963-148702	Sequence 148702,
40	49	28.2	2649	16	US-10-437-963-168995	Sequence 168995,
41	48.5	27.9	107	15	US-10-282-122A-72513	Sequence 72513, A
42	48.5	27.9	1465	15	US-10-282-122A-72528	Sequence 72528, A
43	48	27.6	47	14	US-10-029-386-29475	Sequence 29475, A
44	48	27.6	267	14	US-10-369-493-8315	Sequence 8315, Ap
45	48	27.6	640	16	US-10-437-963-195690	Sequence 195690,

## ALIGNMENTS

RESULT 1  
US-10-092-750-7  
; Sequence 7, Application US/10092750  
; Publication No. US20030032157A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; APPLICANT: Alpin, Julia  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1  
; FILE REFERENCE: 50036/050002  
; CURRENT APPLICATION NUMBER: US/10/092, 750  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: US 60/274,526  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-750-7

Query Match 100.0%; Score 174; DB 14; Length 35;  
Best Local Similarity 100.0%; Pred. No. 3.1e-18;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTDRDPRMFTWAIETIQLRLTGVDKMAEYNTA 35  
DB 1 GTDRDPRMFTWAIETIQLRLTGVDKMAEYNTA 35

RESULT 2  
US-09-925-299-1540  
; Sequence 1540, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:

```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,239
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1540
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-1540

Query Match          44.3%; Score 77; DB 9; Length 57;
Best Local Similarity 55.2%; Pred. No. 0.00098;
Matches 16; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Cy 1 GTRDDMFETMAIEIQLARLVGVNDKM 29
Db 29 GSSQDRMFVTYVLEIEQLARVGVLDKSI 57

RESULT 3
; US-09-925-299-1540
; Sequence 1540, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,239
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1540
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-299-1540

Query Match          44.3%; Score 77; DB 10; Length 57;
Best Local Similarity 55.2%; Pred. No. 0.00098;
Matches 16; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Cy 1 GTRDDMFETMAIEIQLARLVGVNDKM 29
Db 29 GSSQDRMFVTYVLEIEQLARVGVLDKSI 57

RESULT 4
; US-10-424-599-148251
; Sequence 148251, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(55223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 148251
```

```

; LENGTH: 241
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_104893C.1.psp
; US-10-424-599-148251

Query Match          32.8%; Score 57; DB 15; Length 241;
Best Local Similarity 37.0%; Pred. No. 4.7;
Matches 10; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Cy 6 RMEETMAIEIQLARLVGVNDKMAEY 32
Db 60 RSWKSMEMETQSLERKLDINDSMRW 86

RESULT 5
; US-10-369-493-9711
; Sequence 9711, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9711
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Desulfitobacterium hafnense
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(601)
; OTHER INFORMATION: unsure at all Xaa locations
; US-10-369-493-9711

Query Match          32.8%; Score 57; DB 14; Length 601;
Best Local Similarity 41.4%; Pred. No. 14;
Matches 12; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

Cy 3 RQDMFETMAIEIQLARLVGVNDKMAE 31
Db 186 RTDPRYDSXDEIENLKRIPGIGSKWAK 214

RESULT 6
; US-09-815-242-13668
; Sequence 13668, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Mail, Daniel
; APPLICANT: Treawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

```

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13668
; LENGTH: 1460
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-815-242-13668
```

```
Query Match          31.3%; Score 54.5; DB 9; Length 1460;
Best Local Similarity 36.8%; Pred. No. 93;
Matches 14; Conservative 11; Mismatches 8; Indels 5; Gaps 3;
```

```
OY      2 TRQDRM-FETMAIEIQLARLTG--VNDKMAEYTN 35
DB      237 TEENRLVPEGVDFVEQKTR-TGRVLINFKMTDYTS 273
```

```
RESULT 7
US-10-282-122A-73765
; Sequence 73765, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
```

```

; SEQ ID NO 73765
; LENGTH: 1463
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-10-282-122A-73765
```

```
Query Match          31.3%; Score 54.5; DB 15; Length 1463;
Best Local Similarity 36.8%; Pred. No. 93;
Matches 14; Conservative 11; Mismatches 8; Indels 5; Gaps 3;
```

```
OY      2 TRQDRM-FETMAIEIQLARLTG--VNDKMAEYTN 35
DB      237 TEENRLVPEGVDFVEQKTR-TGRVLINFKMTDYTS 273
```

```
RESULT 8
US-09-738-626-4288
; Sequence 4288, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4288
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4288
```

```
Query Match          31.0%; Score 54; DB 9; Length 387;
Best Local Similarity 38.5%; Pred. No. 23;
Matches 15; Conservative 3; Mismatches 11; Indels 10; Gaps 1;

OY      6 RMFETMAIEIQLARLTGVNDKMA-----EYTN 34
DB      14 RMQTSALIELGSLRVLAGAPDSSAAQVADVVGDTN 52

RESULT 9
US-09-815-242-10331
; Sequence 10331, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
```

```
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10331
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10331

Query Match
  31.0%; Score 54; DB 9; Length 1034;
Best Local Similarity 40.7%; Pred. No. 73;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Cy
  1 GTGDDMEETMAIEIQLARLTGVNDK-MAEY 27
Db
  148 GTTQDDISDYVANSVKDTLSRLNGVD 174

RESULT 10
US-09-950-368-1
; Sequence 1, Application US/09950368
; Patent No. US20020061580A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCT
; CURRENT APPLICATION NUMBER: US/09/950,368
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 09/355,166
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Bacillus
US-09-950-368-1

Query Match
  30.2%; Score 52.5; DB 9; Length 657;
Best Local Similarity 42.4%; Pred. No. 71;
Matches 14; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Cy
  2 TRODRMEETMAIEIQLARLTGVNDK-MAEY 33
Db
  365 TKDPRSELXSYIPLGQEKQLTGANDKFEVREHT 397

RESULT 11
US-10-401-437-2
; Sequence 2, Application US/10401437
; Publication No. US20030158070A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC382-US
; CURRENT APPLICATION NUMBER: US/10/401,437
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: PCT/US98/14647
```

```
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: BP 97305237.7
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-401-437-2

Query Match
  30.2%; Score 52.5; DB 14; Length 657;
Best Local Similarity 42.4%; Pred. No. 71;
Matches 14; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Cy
  2 TRODRMEETMAIEIQLARLTGVNDK-MAEY 33
Db
  365 TKDPRSELXSYIPLGQEKQLTGANDKFEVREHT 397

RESULT 12
US-10-402-312-2
; Sequence 2, Application US/10402312
; Publication No. US20030166248A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC382-US
; CURRENT APPLICATION NUMBER: US/10/402,312
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: PCT/US98/14647
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: BP 97305237.7
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-402-312-2

Query Match
  30.2%; Score 52.5; DB 14; Length 657;
Best Local Similarity 42.4%; Pred. No. 71;
Matches 14; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Cy
  2 TRODRMEETMAIEIQLARLTGVNDK-MAEY 33
Db
  365 TKDPRSELXSYIPLGQEKQLTGANDKFEVREHT 397

RESULT 13
US-10-402-067-2
; Sequence 2, Application US/10402067
; Publication No. US20030175892A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC382-US
; CURRENT APPLICATION NUMBER: US/10/402,067
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: PCT/US98/14647
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: BP 97305237.7
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-402-067-2
```

	Query Match	30.2%	Score 52.5	DB 14	Length 657
	Best Local Similarity	42.4%	Pred. No. 71		
Matches	14	Conservative	6	Mismatches 12	Indels 1
QY	2	TRQDRPFETMAIEHQLARLTGVNDK-MAERT	33		
		: : : : : : : : : : : : : : : :			
Db	365	TKPRDPSRLYSLPQGEKQLTGANDKFEVREHT	397		

```

RESULT 14
US-10-401-436-2
; Sequence 2, Application US/10401436
; Publication No. US20030180532A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC382-US
; CURRENT APPLICATION NUMBER: US/10/401,436
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: PCT/US98/14647
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: EP 97305237.7
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
;
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
;
; LENGTH: 657
;
; TYPE: prt
;
; ORGANISM: Bacillus subtilis
;
; US-10-401-436-2

```

```

Query Match          30.2%; Score 52.5; DB 14; Length 657;
Best Local Similarity 42.4%; Pred. No. 71;
Matches 14; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

OY      2 TRQDRMEETMAIEIEQLLARTGVNDK-MAEYT 33
       ||| | | | | | | | | | | | | | | |
Db      365 TKPDRPEELYSIPLGGEKKLTGANDKFVAEHT 397

RESULT_15
US-10-437-963-192198
Sequence 192198, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437, 963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 192198
LENGTH: 345
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_B8447C.1.pep
US-10-437-963-192198

```

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:36:51 ; Search time 7.07562 Seconds  
(without alignments)  
475.942 Million cell updates/sec

Title: US-10-092-750-7

Sequence: 1 GTRQDRMFETMAIEIQLLARLTGVNDKKAETNA 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR\_79:\*  
2: pir1:\*  
3: pir2:\*  
4: pir3:\*  
5: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	35.1	1741	2 T15978	hypothetical prote
2	59	33.9	821	2 S67087	hypothetical prote
3	59	33.9	985	2 A96777	hypothetical prote
4	57.5	33.0	527	2 E87610	alkyl hydroperoxid
5	55	31.6	239	2 F84887	probable cis-Golgi
6	54.5	31.3	1463	2 C95032	hypothetical prote
7	54.5	31.3	1463	2 C97903	hypothetical prote
8	54	31.0	1034	2 D65119	DNA-directed DNA p
9	53	30.5	1140	2 AE0180	acridin resista
10	52.5	30.2	657	2 E70025	probable membrane
11	50	28.7	213	2 E81049	glutaredoxin 2 NMA
12	50	28.7	337	2 B54497	surface membrane p
13	50	28.7	507	2 G70604	DNA gyrase, subunit
14	50	28.7	786	2 A69308	probable acid-CoA
15	49.5	28.4	611	2 A82358	vitamin B12 recept
16	49.5	28.4	1728	2 T17466	riboflavin polyketi
17	49	28.2	513	2 D86273	hypothetical prote
18	49	28.2	513	2 I39534	hypothetical prote
19	49	28.2	649	2 F86740	hypothetical prote
20	49	28.2	752	2 D72726	hypothetical prote
21	49	28.2	778	2 D87965	protein Y633A.6b
22	49	28.2	1133	2 T22976	hypothetical prote
23	49	28.2	1181	2 T01799	hypothetical prote
24	49	28.2	1510	2 D64120	cell division prot
25	48.5	27.9	581	2 S73882	GTP-binding membra
26	48	27.6	322	2 E96639	protein fil9.10 (I
27	48	27.6	322	2 A81074	probable phosphos
28	48	27.6	328	2 A82087	phosphoserine phos
29	48	27.6	430	1 A27655	adenosylhomocyste

30	48	27.6	1186	2 G69708	chromosome segrega
31	48	27.6	1438	2 C99900	DNA polymerase III
32	48	27.6	2388	2 JE0271	beta spectrin, bet
33	48	27.6	4572	2 S57908	hypothetical 527K
34	47.5	27.3	568	2 E70952	probable facB4 pro
35	47.5	27.3	1208	2 E70952	chromosome segrega
36	47	27.0	248	2 T22342	hypothetical prote
37	47	27.0	416	2 D90042	fmha protein limpo
38	47	27.0	474	2 F86819	dipeptidase limpor
39	47	27.0	489	2 T44978	transducer protein
40	47	27.0	489	2 E84304	Hr10 transducer
41	47	27.0	629	2 T24459	hypothetical prote
42	47	27.0	1024	2 AB0347	Acid/acid/acid fam
43	47	27.0	1026	2 AB0771	probable RND-fam1
44	47	27.0	1228	2 G96751	hypothetical prote
45	46.5	26.7	282	2 E70890	hypothetical prote

## ALIGNMENTS

### RESULT 1

115978 hypothetical protein F08P8.4 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 23-Mar-2001

C/Accession: T15978

R/Du: Z.

submitted to the EMBL Data Library, June 1995

A/Description: The sequence of C. elegans cosmid F08P8.

A/Reference number: S61147

A/Accession: T15978

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1741 <DUZ>

A/Cross-references: EMBL:U28991; NID:G861364; PID:G861366; PIDN:AAA68385.1; CESP:F08P8.

A/Experimental source: strain Bristol N2

C/Genetics:

A/Gene: CESP.F08P8.4

A/Introns: 33/1; 64/1; 130/3; 201/2; 251/3; 362/2; 568/2; 628/3; 648/3; 655/3; 669/3; 7/2; 1676/3

Query Match

Best Local Similarity 42.9%; Pred. No. 5.4; Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 8 FETMAIEIQLARLTGVNDKKAETNA 35

DB 1340 KFTVTEIEGLIQLTINDMDVAGA 1367

### RESULT 2

S67087 hypothetical protein YOR195w - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein O4806

C/Species: Saccharomyces cerevisiae

C/Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 09-Jul-2004

C/Accession: S67087

R/Hughes, B.; Pohl, T.M.

submitted to the Protein Sequence Database, July 1996

A/Reference number: S66685

A/Accession: S67087

A/Molecule type: DNA

A/Residues: 1-821 <HUG>

A/Cross-references: UNIPROT:Q08581; EMBL:Z75103; NID:G1420464; PID:e252389; PID:G142046

A/Experimental source: strain S288C

C/Genetics:

A/Gene: SGD:SLK19; MIPS:YOR195w

A/Cross-references: SGD:S0005721

A/Map position: 15R

Query Match

Best Local Similarity 33.3%; Score 59; DB 2; Length 821; Pred. No. 4.4;

## RESULT 5

RESULT 7  
C97903  
DNA-directed DNA polymerase (EC 2.7.7.7) III, alpha chain [imported] - Streptococcus pneumoniae  
C/Species: Streptococcus pneumoniae  
C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C/Accession: C97903  
R/Hoskins, J.A.; Albom Jr., W.; Arnold, J.; Blaszczak, L.; Burgess, S.; Dehoff, B.S.;  
e, R.; Ielanc, D.U.; Lee, L.N.; Lefkowitz, E.J.; Lu, Y.; Matsushita, P.; McAnen, S.;  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A/Authors: Yang, Y., Young-Bellido M.; Zhao, G.; Zock, C.; Baltz, R.H.; Jaskunas, S.R  
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain K6.



A:Reference number: A97872; MUID:214229245; PMID:11544224  
A:Accession: C97903  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-1463 <KUR>  
A:Cross-references: UNIPROT:O8DPA5; GB:AEO07317; PIDDN:AAK9055.1; PID:G15457800; GSPDB:G:  
C:Genetics:  
A:Gene: polC  
C:Superfamily: DNA-directed DNA polymerase III alpha chain polC  
C:Keywords: nucleotidyltransferase

Query Match                      31.3%; Score 54.5; DB 2; Length 1463;  
Best Local Similarity        36.8%; Pred. No.36;  
Matches     14; Conservative    11; Mismatches    8; Indels        5; Gaps        3;

OY                      2 TRODRM-FETMAIEEQLARLTG---VNDCKAEYTA 35  
                        ::.:|||::|||::|||::|||::  
DB                      237 TEENRLVESGVFDEQCKTR-TGRVLINPOTHTI 273

**RESULT 8**

D65119  
acriflavin resistance protein acrF - Escherichia coli (strain K-12)

N/Alternate names: envD protein  
N/Species: Escherichia coli  
C/Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
A/Acessionion: D65119; S18537  
R/Baltnner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A./Rose, D.J.; Mau, B.; Shaq, Y.  
Science 277, 1453-1462, 1997

A>Title: The complete genome sequence of Escherichia coli K-12.  
A/Reference number: A64720; MUID:97426617; PMID:9278503

A/Acessionion: D65119  
A>Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA  
A/Residues: 1-1034 <BLAT>

A/Cross-references: UNIPROT:P24181; GB:AEO00405; GB:U00096; NID:g1789659; PIDN:AACT6296.  
A/Experimental source: Strain K-12, Substrains MG1655  
R/Klein, J.R.; Henrich, B.; Plapp, R.  
Mol. Gen. Genet. 230, 230-240, 1991

A>Title: Molecular analysis and nucleotide sequence of the envCD operon of Escherichia c  
A/Reference number: S18536; MUID:92079901; PMID:11720861

A/Acessionion: S18537  
A>Status: Preliminary  
A/Molecule type: DNA  
A/Residues: 69-222, 'N', 224-283, 'T', 285, 'PP', 288-337, 'KCKNAVRSHVAGVPEDGDSVLAEYASNADPHHCGR  
S-1034 <XLE>

A/Cross-references: EMBL:X57948; NID:g510827; PIDN:CMA41017.1; PID:g510830

A/Experimental source: strain K-12

C/Genetics:  
A/Gene: acrF; envD  
C/Superfamily: acriflavin resistance protein  
C/Keywords: cell division; transmembrane protein; transport protein  
F/1-28/Domain: transmembrane #status predicted <TM1>  
F/343-359/Domain: transmembrane #status predicted <TM2>  
F/370-386/Domain: transmembrane #status predicted <TM3>  
F/397-413/Domain: transmembrane #status predicted <TM4>  
F/442-458/Domain: transmembrane #status predicted <TM5>  
F/472-488/Domain: transmembrane #status predicted <TM6>  
F/540-556/Domain: transmembrane #status predicted <TM7>  
F/574-590/Domain: transmembrane #status predicted <TM8>  
F/698-914/Domain: transmembrane #status predicted <TM9>  
F/974-990/Domain: transmembrane #status predicted <TM10>  
F/1012-1028/Domain: transmembrane #status predicted <TM11>

Query Match 31.0%; Score 54; DB 2; Length 1034;  
Best Local Similarity 40.7%; Pred. No. 29;  
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

CY 1 GTGGRRFFETMAIEIQLLRITGVND 27  
||| ||| : : : : : : : : : :  
148 GTTGDDISDVVASNVKDTLSRLNGVED 174

```

RESULT 9
AE0180
probable membrane protein YP01482 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #ext_change 09-Jul-2004
C:Accession: AE0180
R:Packhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.;
deno-Tarrag, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall
Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11585360
A:Accession: AE0180
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1140 <KTR>
A:Cross-references: UNIPROT:Q8ZG38; GB:AL590842; PIDD:CA0304.1; PID:G15979523; GSPDB
A:Genetics:
A:Gene: YP01482

Query Match      30.5%; Score 53; DB 2; Length 1140;
Best local Similarity 31.4%; Pred. No. 4;
Matches 11; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY      1  GTGGDMFEIWAIEIQLIARLTGVDKAEYTN 35
      |||:::|||||:::|
DB      789  GDAQRRLSDDSLSLQTLITRTVRLKIQGVNA 823

```

RESULT 10  
E70025  
Probable acylaminoadyl-peptidase (EC 3.4.19.1) yuxL - *Bacillus subtilis*  
CDSpecies: *Bacillus subtilis*  
CDate: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
R/Accession: E70025, D25364  
R/Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Allon, G.; Azevedo, V.; Bertet  
C.; Bron, S.; Broiliet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
C.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Erington, J.; Fabre, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fura, S.; Galizzi, A.; Gall  
lech, U.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hu  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kunitz, K.; Lapidus, A.; Lard  
A:Authors: Lauber, U.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Ma  
Y, M.; Ogawa, K.; Ogatawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.;  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Sc  
A:Authors: Schlecht, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.;  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpest, P.; Tognoni, A.; Tosato, V.; Uchiyama  
T.; Winers, P.; Wipit, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida,  
A:Authors: Yoshikawa, H.F.; Zinnstein, E.; Yoshikawa, H.; Danchin, A.  
A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*  
A:Reference number: A69580; MUID:96044033; PMID:9384377  
A:Accession: E70025  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Mcleuc type: DNA  
A:Residues: 1-657 <KUN>  
A:Cross-references: UNPROT:P38839; GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAH521  
A:Experimental source: strain 168  
R:Parrot, C.  
EMBO J. 5, 3013-3019, 1986  
A>Title: Evolution of biosynthetic pathways: a common ancestor for threonine synthase,  
A:Reference number: A91055; MUID:870802861; PMID:3098560  
A:Accession: D25364  
A:Mcleuc type: DNA  
A:Residues: 19PSVQGRRAAPFAGDKGLLEITNCKGGEYE, '3-17','S', '19-24 <PAR>  
A:Cross-references: GB:A04603; NID:ig40210; PIDN:CA28272.1; PID:940213  
A:Note: hypothetical fragment that is probably a mistranslation of a portion of the yuxL  
C:Genetics:  
A:Gene: yuxL  
C:Keywords: hydrolase; omega peptidase  
F:518,599,631/Active site: Ser, Asp, His #status predicted  
Query Match 30.2%; Score 52.5; DB 2; Length 657;

Best Local Similarity 42.4%; Pred. No. 28;  
Matches 14; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

QY 2 TRODRMFETMAIEIQLARLTGVNDK-MAEYT 33  
DB 365 TKPDRPSELVSIPIQGEKQLTGANDKVFREHT 397

## RESULT 11

E81049  
glutaredoxin 2 NMA1990 [imported] - Neisseria meningitidis (strain MC58 serogroup B, ser  
C/Species: Neisseria meningitidis  
C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C/Accession: E81049; C81828  
R/Retelien, H.; Sanders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
R/Hecky, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Hickey, H.; Qim, H.; Vamathevan, J.; Gill, J.; Scariatto, V.; Maignan, V.; Pizzi, M.  
Science 287, 1809-1815, 2000  
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A/Reference number: A81000; PMID:20175755; PMID:10710307  
A/Accession: E81049

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-213 <TEP>  
A/Cross-references: UNIPROT:Q9JH71; GB:AEO02523; GB:AEO02098; NID:g7226980; PIDN:AAE4207  
R/Experimental source: serogroup B, strain MC58  
R/Pathhill, J.; Achtmann, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
R/Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Murgall, K.; Quail, M.A.; Rajandream,  
Nature 406, 502-506, 2000  
A/Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A/Reference number: A81775; PMID:20225556; PMID:10761919  
A/Accession: C81828  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-213 <PAR>  
A/Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAH85210.1; PID:g738062  
A/Experimental source: serogroup A, strain Z2491  
A/Genetics:  
A/Genes: grxB; NMB1734; NMA1990

Query Match 28.7%; Score 50; DB 2; Length 213;  
Best Local Similarity 34.5%; Pred. No. 18;  
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 3 RODRMFETMAIEIQLARLTGVNDKMAE 31  
DB 71 REDRLKDEVRPEIRAWLDKVGNDKLYQ 99

## RESULT 12

B54497  
surface membrane protein p26 precursor - Babesia rodhaini (strain Antwerp)  
C/Species: Babesia rodhaini  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: B54497  
R/Smay, D.; Smith, M.A.  
Mol. Biochem. Parasitol. 27, 303-312, 1988  
A/Title: Sequence homology of surface membrane proteins of Babesia rodhaini.  
A/Reference number: A54497; PMID:88143000; PMID:2893577  
A/Accession: B54497  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-337 <SNA>  
A/Cross-references: UNIPROT:Q17278; GB:M19145; NID:g155924; PIDN:AAA27814.1; PID:g155925  
C/Superfamily: Babesia surface membrane protein p26

Query Match 28.7%; Score 50; DB 2; Length 337;  
Best Local Similarity 35.5%; Pred. No. 30;  
Matches 11; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 2 TRODRMFETMAIEIQLARLTGVNDKMAEY 32

DB 236 TEAKTFDELKNEIEDLVTKNSEKDKAAEH 266

## RESULT 13

G70604  
probable acid-CoA ligase (EC 6.2.1.-) fadJ3 - Mycobacterium tuberculosis (strain H37Rv)  
C/Species: Mycobacterium tuberculosis  
C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C/Accession: G70604  
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,  
R.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; PMID:98295987; PMID:9634230  
A/Accession: G70604  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-507 <COL>  
A/Cross-references: UNIPROT:P66843; GB:I292774; GB:AL123456; NID:g3261729; PIDN:CAH07136  
A/Experimental source: strain H37Rv  
C/Genetics:  
A/Genes: fadJ3  
C/Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology  
C/Keywords: acid-thiol ligase  
F/52-503/Domain: acetate-CoA ligase homology <ACLD>

Query Match 28.7%; Score 50; DB 2; Length 507;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 8 FERMAIEIQLARLTGVND 27  
DB 416 FNVYPAEVRQVLAEMDGVAD 435

## RESULT 14

A69308  
DNA gyrase, subunit A (gyrA) homolog - Archaeoglobus fulgidus  
C/Species: Archaeoglobus fulgidus  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C/Accession: A69308  
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
R/Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A/Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Atliach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
A/Reference number: A69250; PMID:98049343; PMID:9389475  
A/Accession: A69308  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-786 <KLE>  
A/Cross-references: UNIPROT:Q29785; GB:AEO01072; GB:AEO00782; NID:g2689395; PIDN:AAH907  
C/Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase  
F/1-232/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology <T4>

Query Match 28.7%; Score 50; DB 2; Length 786;  
Best Local Similarity 37.9%; Pred. No. 77;  
Matches 11; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 4 QDRMFETMAIEIQLARLTGVNDKMAEY 32  
DB 413 QMRLOKLTAMIDALLKEVSEKXIAEY 441

## RESULT 15

A82338  
vibrio B12 receptor VC0156 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C/Species: Vibrio cholerae  
C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C:Accession: A82358  
 R:Heidelberg, U.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Yamashvyan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.  
 L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: A82358  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-611 <HEI>  
 A:Cross-references: UNIPROT:Q9KV19; GB:AE004105; GB:AE003852; NID:99654551; PIDN:AAF333  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VC0156  
 A:Map position: 1  
 C:Superfamily: vitamin B12 receptor; tonB-dependent receptor amino-terminal homology; td

Query Match 28.4%; Score 49.5; DB 2; Length 611;  
 Best Local Similarity 54.2%; Pred. No. 68;  
 Matches 13; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 2 TRQDRMFETMAIEQLRLRGV 25  
 ||||| : ||| : |||||  
 DB 53 TRQD-IEQIQAKTLPRLRLRGV 75

Search completed: November 10, 2004, 12:29:08  
 Job time : 9.07562 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:27:01 ; Search time 38.2948 Seconds  
(without alignments)  
525.871 Million cell updates/sec

Title: US-10-092-750-7  
Perfect score: 174  
Sequence: 1 GTRODMFETMAIEEQIARLTGVNDKAEYTNVA 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	161	92.5	175	2	Q96Q19
2	161	92.5	250	1	Q96Q19
3	161	92.5	250	1	Q96Q19
4	156	89.7	250	1	Q96Q19
5	156	89.7	250	1	Q96Q19
6	156	89.7	250	1	Q96Q19
7	156	89.7	250	1	Q96Q19
8	142	81.6	251	2	Q96Q19
9	109	62.6	179	2	Q96Q19
10	109	62.6	179	2	Q96Q19
11	109	62.6	179	2	Q96Q19
12	109	62.6	179	2	Q96Q19
13	86	49.4	241	2	Q96Q19
14	63	36.2	152	2	Q96Q19
15	63	36.2	152	2	Q96Q19
16	60	34.5	232	1	Q96Q19
17	59	33.9	325	2	Q96Q19
18	59	33.9	325	2	Q96Q19
19	59	33.9	325	2	Q96Q19
20	59	33.9	325	2	Q96Q19
21	57.5	33.0	527	2	Q96Q19
22	57	32.8	339	2	Q96Q19
23	56	32.2	413	1	Q96Q19
24	55	31.6	239	1	Q96Q19
25	55	31.6	436	1	Q96Q19
26	55	31.6	436	1	Q96Q19
27	55	31.6	436	1	Q96Q19
28	55	31.6	436	1	Q96Q19
29	55	31.6	436	1	Q96Q19
30	54.5	31.3	761	2	Q96Q19
31	54.5	31.3	1463	1	Q96Q19

32	54.5	31.3	1463	1	DPO3 STRR6	Q8d7a5 streptococc
33	54	31.0	203	2	Q7V4S3	Q7V4S3 prochloroc
34	54	31.0	387	2	Q8NSG4	Q8NSG4 corynebacte
35	54	31.0	399	2	Q6W768	Q6W768 corynebacte
36	54	31.0	399	2	Q6W768	Q6W768 corynebacte
37	54	31.0	1034	1	ACRP ECOLI	P24181 escherichia
38	54	31.0	1034	2	Q8FDZ5	Q8FDZ5 escherichia
39	54	31.0	1213	2	Q6BJL1	Q6BJL1 debaryomyce
40	53	30.5	194	2	Q7MT00	Q7MT00 porphyromon
41	53	30.5	396	2	Q82UT6	Q82UT6 nitrosomonas
42	53	30.5	1013	2	Q7C210	Q7C210 shigella fl
43	53	30.5	1013	2	Q812L9	Q812L9 shigella fl
44	53	30.5	1140	2	Q8ZG38	Q8ZG38 yersinia pe
45	53	30.5	1140	2	AA561613	AA561613 yersinia

ALIGNMENTS

RESULT 1	Q96Q19	PRELIMINARY;	PRT;	175 AA.
ID	Q96Q19			
AC	Q96Q19			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	GOSRL protein.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euleleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RX	MEDLINE=2388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Steinem C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Ueding T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raba S.S., Longellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.W., Gay L.V., Hulyk S.W.,			
RA	Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smalios D.E., Schmechel A., Schein J.E.,			
RA	Jones S.V., Maitra M.A.,			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."/			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Strausberg R.,			
RL	Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC012620; AA012620.1; -			
DR	GO; GO:0016021; C:intracellular membrane; IEA.			
DR	GO; GO:0006886; P:intracellular protein transport; IEA.			
DR	InterPro; IPR007705; V-SNARE.			
DR	Pfam; PF05008; V-SNARE.1.			
DR	SEQUENCE 175 AA; 20068 MW; B11D29D4B05235B0 CRC64;			
QY	Query Match	92.5%;	Score 161;	DB 2; Length 175;
	Best Local Similarity	91.4%;	Pred. No. 3.5e-14;	
	Matches	32;	Conservative	2; Mismatches 1; Indels 0; Gaps 0;
	1 GTRODMFETMAIEEQIARLTGVNDKAEYTNVA 35			
	:			

Db 60 GSSQDRMFETMAIEIQLLRLTGVDKMAEYTN 94

RESULT 2

GS28\_CRIGR STANDARD; PRT; 250 AA.

AC 008522;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE 28 kDa Golgi SNARE protein (Golgi SNARE receptor complex member 1) (28 kDa cis-Golgi SNARE p28) (GOS-28).

OS Name=GOSR1; Synonyms=GS28;

OC Cricetus griseus (Chinese hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

OC Cricetus.

OX NCBI\_Taxid=10029;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=9622356; PubMed=8636227;

RA Nagahama M., Orci L., Ravazzola M., Amherdt M., Lacomis L., Tempst P., Rothman J.E., Soellner T.H.;

RT "A V-SNARE implicated in intra-Golgi transport.";

RL J. Cell Biol. 133:507-516(1996).

CC -1- FUNCTION: Involved in transport from the ER to the Golgi apparatus as well as in intra-Golgi transport.

CC -1- SUBCELLULAR LOCATION: Type IV membrane protein. Enriched on vesicular components at the terminal rims of the Golgi.

CC -1- SIMILARITY: Belongs to the GOSR1 family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U69841; AAB51019.1; -

DR GO: GO:0000138; C:Golgi trans cisterna; IDA.

DR GO: GO:0005484; F:SNARE receptor activity; NAS.

DR GO: GO:0006888; P:ER to Golgi transport; IDA.

DR GO: GO:0006891; P:intra-Golgi transport; IDA.

DR InterPro: IPR007705; V-SNARE.

DR Pfam: PF05008; V-SNARE; 1.

KW Coiled coil; Golgi stack; Protein transport; Transmembrane; Transport.

FT DOMAIN 1 229 Cytoplasmic (Potential).

FT TRANSMEM 230 250 Anchor for type IV membrane protein (Potential).

FT FT 9 30 Coiled coil (Potential).

FT FT 68 95 Coiled coil (Potential).

SO SEQUENCE 250 AA; 28510 MW; BDDA4DCB808068DD CRC64;

Query Match 92.5%; Score 161; DB 1; Length 250;

Best Local Similarity 91.4%; Pred. No. 5.2e-14;

Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTRQDRMFETMAIEIQLLRLTGVDKMAEYTN 35

Db 60 GSSQDRMFETMAIEIQLLRLTGVDKMAEYTN 94

RESULT 3

GS28\_HUMAN STANDARD; PRT; 250 AA.

ID GS28; 075392;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE 28 kDa Golgi SNARE protein (Golgi SNARE receptor complex member 1) (28 kDa cis-Golgi SNARE p28) (GOS-28).

OS Name=GOSR1; Synonyms=GS28;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Umbilical cord blood;

RX MEDLINE=9831851; PubMed=9653160;

RA Mao M., Fu G., Wu J.-S., Zhang Q.-H., Zhou J., Kan L.-X., Huang Q.-H., He K.-L., Gu B.-W., Han Z.-G., Shen Y., Gu J., Yu Y.-P., Xu S.-H.;

RA Wang Y.-X., Chen S.-D., Chen Z.;

RT "Identification of genes expressed in human CD34(+) hematopoietic stem/progenitor cells by expressed sequence tags and efficient full-length cDNA cloning.";

RL Proc. Natl. Acad. Sci. U.S.A. 95:8175-8180(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Colon carcinoma;

RX MEDLINE=99216422; PubMed=10198168;

RA Bui T.D., Levy E.R., Subramaniam V.N., Lowe S.L., Hong W.;

RT "CDNA characterization and chromosomal mapping of human Golgi SNARE G527 and GS28 to chromosome 17.";

RL Genomics 57:285-288(1999).

CC -1- FUNCTION: Involved in transport from the ER to the Golgi apparatus as well as in intra-Golgi transport (By similarity).

CC -1- SUBUNIT: Interacts with GABARAPL2 (By similarity).

CC -1- SUBCELLULAR LOCATION: Type IV membrane protein. Enriched on vesicular components at the terminal rims of the Golgi (By similarity).

CC -1- SIMILARITY: Belongs to the GOSR1 family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AF047438; AAC39889.1; -

DR EMBL: AF073926; AAD12945.1; -

DR Genew; HGNC:4430; GOSR1.

DR MIM; 604026; -

DR GO: GO:0005485; F:V-SNARE activity; TAS.

DR GO: GO:0006891; P:intra-Golgi transport; TAS.

DR InterPro: IPR007705; V-SNARE.

DR Pfam: PF05008; V-SNARE; 1.

KW Coiled coil; Golgi stack; Protein transport; Transmembrane; Transport.

FT DOMAIN 1 229 Cytoplasmic (Potential).

FT TRANSMEM 230 250 Anchor for type IV membrane protein (Potential).

FT FT 9 30 Coiled coil (Potential).

FT FT 68 95 Coiled coil (Potential).

FT FT 11 11 D -> GEARPPD (in Ref. 1).

FT FT 49 50 Missing (in Ref. 1).

FT FT 156 157 NN -> TT (in Ref. 1).

FT FT 159 159 R -> G (in Ref. 1).

FT FT 170 170 L -> F (in Ref. 1).

SO SEQUENCE 250 AA; 28612 MW; EEDA4291436152F6 CRC64;

Query Match 92.5%; Score 161; DB 1; Length 250;

Best Local Similarity 91.4%; Pred. No. 5.2e-14;

Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTRQDRMFETMAIEIQLLRLTGVDKMAEYTN 35

Db 60 GSSQDRMFETMAIEIQLLRLTGVDKMAEYTN 94

RESULT 4

GS28\_MOUSE STANDARD; PRT; 250 AA.

ID GS28\_MOUSE

AC 088630;

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 28 kDa Golgi SNARE protein (Golgi SNAP receptor complex member 1) (28
DE kDa cis-Golgi SNARE p28) (GOS-28).
GN Name=Gosrl; Synonyms=Gsz28;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Bui T.D., Subramaniam V.N., Hong W.;
RU Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in transport from the ER to the Golgi apparatus
CC as well as in intra-Golgi transport (By similarity).
CC -1- SUBUNIT: Interacts with GABARAPL2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Type IV membrane protein. Enriched on
CC vesicular components at the terminal rims of the Golgi (By
CC similarity).
CC -1- SIMILARITY: Belongs to the GOSR1 family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF079901; AAC32189.1; -
DR MGI; MGI:1858260; Gosrl.
DR GO; GO:0000139; C:Golgi membrane; IDA.
DR InterPro; IPR007705; V-SNARE.
DR Pfam; PF05008; V-SNARE.1.
KM Coiled coil; Golgi stack; Protein transport; Transmembrane; Transport.
FT DOMAIN 1 229 Cytoplasmic (Potential).
FT TRANSMEM 230 250 Anchor for type IV membrane protein
FT FT (Potential).
FT DOMAIN 9 27 Coiled coil (Potential).
FT FT 72 93 Coiled coil (Potential).
SQ SEQUENCE 250 AA; 28429 MW; 9B3D83FB05249B9D CRC64;

Query Match 89.7%; Score 156; DB 1; Length 250;
Best Local Similarity 88.6%; Pred. No. 2,6e-13;
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTRODMFETMAIEIQLARLTGVNDKMAEYTN 35
DB 60 GSSQDMFETMAIEIQLARLTGVNDKMAEYTHS 94

RESULT 5
GS28 RAT STANDARD; PRT; 250 AA.
ID GS28 RAT STANDARD; PRT; 250 AA.
AC Q62931.1
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 28 kDa Golgi SNARE protein (Golgi SNAP receptor complex member 1) (28
DE kDa cis-Golgi SNARE p28) (GOS-28).
GN Name=Gosrl; Synonyms=Gsz28;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-25; 50-53; 76-86; 113-121;
RA 127-133; 145-149; 160-165; 176-187 AND 212-225.
RC TISSUE=Brain;
RX MEDLINE=96216289; Pubmed=6638159;
RA Subramaniam V.N., Peter F., Philip R., Wong S.H., Hong W.;
RA "GS28", a 28-kilodalton Golgi SNARE that participates in ER-Golgi

```

```

RT transport.";
RL Science 272:1161-1163(1996).
RN [2]
RP SEQUENCE OF 18-26 AND 66-80.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=97248495; Pubmed=90947723;
RA Hay U.C., Chao D.S., Kuo C.S., Scheller R.H.;
RT "Protein interactions regulating vesicle transport between the
RT endoplasmic reticulum and Golgi apparatus in mammalian cells";
RL Cell 89:149-158(1997).
CC -1- FUNCTION: Involved in transport from the ER to the Golgi apparatus
CC as well as in intra-Golgi transport.
CC -1- SUBUNIT: Interacts with GABARAPL2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Type IV membrane protein. Enriched on
CC vesicular components at the terminal rims of the Golgi.
CC -1- SIMILARITY: Belongs to the GOSR1 family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U49099; AAC52597.1; -
DR RGD; 71093; Gosrl.
DR InterPro; IPR007705; V-SNARE.
DR Pfam; PF05008; V-SNARE.1.
KM Coiled coil; Direct protein sequencing; Golgi stack;
KM Protein transport; Transmembrane; Transport.
FT DOMAIN 1 229 Cytoplasmic (Potential).
FT TRANSMEM 230 250 Anchor for type IV membrane protein
FT FT (Potential).
FT DOMAIN 9 30 Coiled coil (Potential).
FT FT 68 95 Coiled coil (Potential).
SQ SEQUENCE 250 AA; 28533 MW; 0F85EDCE7ADA3E2 CRC64;

Query Match 89.7%; Score 156; DB 1; Length 250;
Best Local Similarity 88.6%; Pred. No. 2,6e-13;
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTRODMFETMAIEIQLARLTGVNDKMAEYTN 35
DB 60 GSSQDMFETMAIEIQLARLTGVNDKMAEYTHS 94

RESULT 6
Q80UDU PRELIMINARY; PRT; 250 AA.
ID Q80UDU PRELIMINARY; PRT; 250 AA.
AC Q80UDU;
DT 01-JUN-2003 (TRENDEL. 24, Created)
DT 01-JUN-2003 (TRENDEL. 24, Last sequence update)
DT 01-MAR-2004 (TRENDEL. 26, Last annotation update)
DE Golgi SNAP receptor complex member 1.
GN Name=Gosrl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=129/Sv x 129/Sv-CP; TISSUE=ES Cell;
RX MEDLINE=22388257; Pubmed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buettner K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.U., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Maria M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.",
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RT
RL
RN
RP SEQUENCE FROM N.A.

SQ SEQUENCE 250 AA; 28431 MW; BC0BC03C9C23083 CRC64;

Query Match      89.7%; Score 156; DB 2; Length 250;
Best Local Similarity 88.6%; Pred. No. 2.6e-13;
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTRQDMEFMAIEIQLARLTGVNDKMAEYTN 35
DB 60 GSSQDRMEFMAIEIQLARLTGVNDKMAEYTH 94

RESULT 7
Q91VU9 PRELIMINARY; PRT; 250 AA.
AC Q91VU9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Golgi SNAP receptor complex member 1.
GN Name=Gosr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;

RP SEQUENCE FROM N.A.
RC STRAIN=Mix FVB/N;
RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshitsuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Maria M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.",
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RT
RL
RN
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Mix FVB/N;
RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RA Strausberg R.;
RA Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC008542; AA08542.1;
DR MGD; MG1:1858260; Gosr1.
DR GO; GO:0000139; C:Golgi membrane; IDA.
DR InterPro; IPR007705; V-SNARE.
DR Pfam; PF05008; V-SNARE; 1.
DR Receptor.
SQ SEQUENCE 250 AA; 28489 MW; ACF7PF5CCA0C2D52 CRC64;

Query Match      89.7%; Score 156; DB 2; Length 250;
Best Local Similarity 88.6%; Pred. No. 2.6e-13;
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTRQDMEFMAIEIQLARLTGVNDKMAEYTN 35
DB 60 GSSQDRMEFMAIEIQLARLTGVNDKMAEYTH 94

RESULT 8
Q7ZX43 PRELIMINARY; PRT; 251 AA.
AC Q7ZX43;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gosr1 protein (Fragment).
GN Name=Gosr1;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_Taxid=6355;

RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshitsuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Maria M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.",
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RT
RL
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson F.,
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RA RT initiative.",
RA Dev. Dyn. 225:384-391(2002).
RT
RL
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.,
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

```



DR EMBL; BC045246; AAH45246.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0006886; P:intracellular protein transport; IEA.  
 DR InterPro; IPR007705; V-SNARE.  
 DR Pfam; PF05008; V-SNARE; 1.  
 FT NON TER 1  
 SQ SEQUENCE 251 AA; 28508 MW; 3CB18D70D3834E12 CRC64;

Query Match 81.6%; Score 142; DB 2; Length 251;  
 Best Local Similarity 79.4%; Pred. No. 2.3e-11;  
 Matches 27; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GTRDDMPETMAIEIEQLARLTGVNDKMAEYTN 34  
 Db 61 GSSQDHMETVAVEIEQLAKLTGVNDKMAEYSS 94

RESULT 9  
 ID Q6UIP9 PRELIMINARY; PRT; 179 AA.  
 AC Q6UIP9;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Golgi SNAP receptor complex member 1 (Fragment).  
 GN Name=GOSRI;  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopitheciinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22946167; PubMed=14557539;  
 RA Geschwind D.H., Lockhart D.J., Preuss T.M., Barlow C.;  
 RT "Elevated gene expression levels distinguish human from non-human  
 RT primate brains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:13030-13035(2003).  
 DR EMBL; AY369815; AAR11246.1; -.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR007705; V-SNARE.  
 DR Pfam; PF05008; V-SNARE; 1.  
 KW Receptor.  
 FT NON TER 1  
 SQ SEQUENCE 179 AA; 20502 MW; 27271BB51BE5B282 CRC64;

Query Match 62.6%; Score 109; DB 2; Length 179;  
 Best Local Similarity 95.7%; Pred. No. 6.3e-07;  
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 13 IEIEQLARLTGVNDKMAEYTN 35  
 Db 1 IEIEQLARLTGVNDKMAEYTN 23

RESULT 10  
 ID Q6UIQ0 PRELIMINARY; PRT; 179 AA.  
 AC Q6UIQ0;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Golgi SNAP receptor complex member 1 (Fragment).  
 GN Name=GOSRI;  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22946167; PubMed=14557539;  
 RA Geschwind D.H., Lockhart D.J., Preuss T.M., Barlow C.;  
 RT "Elevated gene expression levels distinguish human from non-human  
 RT primate brains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:13030-13035(2003).  
 DR EMBL; AY369815; AAR11246.1; -.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR007705; V-SNARE.  
 DR Pfam; PF05008; V-SNARE; 1.  
 KW Receptor.  
 FT NON TER 1  
 SQ SEQUENCE 179 AA; 20499 MW; 273AB0DB095B282 CRC64;

RA Geschwind D.H., Lockhart D.J., Preuss T.M., Barlow C.;  
 RT "Elevated gene expression levels distinguish human from non-human  
 RT primate brains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:13030-13035(2003).  
 DR EMBL; AY369814; AAR11245.1; -.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR007705; V-SNARE.  
 DR Pfam; PF05008; V-SNARE; 1.  
 KW Receptor.  
 FT NON TER 1  
 SQ SEQUENCE 179 AA; 20499 MW; 273AB0DB095B282 CRC64;

Query Match 62.6%; Score 109; DB 2; Length 179;  
 Best Local Similarity 95.7%; Pred. No. 6.3e-07;  
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 13 IEIEQLARLTGVNDKMAEYTN 35  
 Db 1 IEIEQLARLTGVNDKMAEYTN 23

RESULT 11  
 ID AAR11245 PRELIMINARY; PRT; 179 AA.  
 AC AAR11245;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DE 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE Golgi SNAP receptor complex member 1 (Fragment).  
 GN GOSRI.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22946167; PubMed=14557539;  
 RA Geschwind D.H., Lockhart D.J., Preuss T.M., Barlow C.;  
 RT "Elevated gene expression levels distinguish human from non-human  
 RT primate brains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:13030-13035(2003).  
 DR EMBL; AY369814; AAR11245.1; -.  
 KW Receptor.  
 FT NON TER 1  
 SQ SEQUENCE 179 AA; 20499 MW; 273AB0DB095B282 CRC64;

Query Match 62.6%; Score 109; DB 2; Length 179;  
 Best Local Similarity 95.7%; Pred. No. 6.3e-07;  
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 13 IEIEQLARLTGVNDKMAEYTN 35  
 Db 1 IEIEQLARLTGVNDKMAEYTN 23

RESULT 12  
 ID AAR11246 PRELIMINARY; PRT; 179 AA.  
 AC AAR11246;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DE 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE Golgi SNAP receptor complex member 1 (Fragment).  
 GN GOSRI.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopitheciinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22946167; PubMed=14557539;  
 RA Geschwind D.H., Lockhart D.J., Preuss T.M., Barlow C.;  
 RT "Elevated gene expression levels distinguish human from non-human  
 RT primate brains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:13030-13035(2003).  
 DR EMBL; AY369815; AAR11246.1; -.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR007705; V-SNARE.  
 DR Pfam; PF05008; V-SNARE; 1.  
 KW Receptor.  
 FT NON TER 1  
 SQ SEQUENCE 179 AA; 20499 MW; 273AB0DB095B282 CRC64;

RA Caceres M., Lachner J., Zapala M.A., Redmond J.C., Kudo L.,  
 RA Geachind D.H., Lockhart D.J., Preuss T.M., Bartlow C.,  
 RT "Elevated gene expression levels distinguish human from non-human  
 RT primate brains";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:13030-13035(2003).  
 DR EMBL; AY358915; AAR1246.1; -  
 KW Receptor.  
 SQ SEQUENCE 179 AA; 20502 MW; 27271BB51B5EB282 CRC64;

Query Match 62.6%; Score 109; DB 2; Length 179;  
 Best Local Similarity 95.7%; Pred. No. 6.3e-07;  
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 IEIQLARLTGVNDKMAEYTN 35  
 DB 1 IEIQLARLTGVNDKMAEYTN 23

## RESULT 13

ID Q7QCS1 PRELIMINARY; PRT; 241 AA.  
 AC Q7QCS1;  
 DT 01-MAR-2004 (TREMBLrel. 26, Created)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE AGC1417 (fragment).  
 GN Name=agc5118; ORFNames=ENSANG00000019523;  
 OS Anopheles gambiae scr. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pezomyzeta;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC Preliminary data.  
 DR EMBL; AAB0100859; EAA08130.1; -  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0006886; P:intracellular protein transport; IEA.  
 DR InterPro; IPR007705; V-SNARE.  
 DR Pfam; PF05008; V-SNARE; 1.  
 FT NON TER  
 SQ SEQUENCE 241 AA; 27244 MW; 01B561E610B41829 CRC64;

Query Match 49.4%; Score 86; DB 2; Length 241;  
 Best Local Similarity 45.2%; Pred. No. 0.0014;  
 Matches 14; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

QY 5 DMFETMAIEIQLARLTGVNDKMAEYTN 35  
 DB 64 DHVFESLSLEIQMDKLSNINKESEIPNS 94

RESULT 14  
 ID Q86E58 PRELIMINARY; PRT; 152 AA.  
 AC Q86E58;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Clone Z2D991 mRNA sequence.  
 OS Schistosoma japonicum (Blood fluke).  
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiida;  
 OC Schistosomatidae; Schistosomatidae; Schistosoma.  
 NCBI\_TaxID=6182;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22879925; PubMed=12973349;  
 RA Hu W., Yan Q., Shen D.K., Liu F., Zhu Z.D., Song H.D., Xu X.R.,

RA Wang Z.J., Rong Y.P., Zeng L.C., Wu J., Zhang X., Wang J.J., Xu X.N.,  
 RA Xue S.Y., Fu G., Zhang X.L., Wang Z.Q., Brindley P.J., McKanus D.P.,  
 RA "Evolutionary and biomedical implications of a Schistosoma japonicum  
 RT complementary DNA resource";  
 RL Nat. Genet. 35:139-147(2003).  
 DR EMBL; AY223362; AAP06399.1; -  
 SQ SEQUENCE 152 AA; 17225 MW; D449AF8FC960A9P CRC64;

Query Match 36.2%; Score 63; DB 2; Length 152;  
 Best Local Similarity 46.4%; Pred. No. 1.4;  
 Matches 13; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 5 DMFETMAIEIQLARLTGVNDKMAEY 32  
 DB 65 DSNFSVMCEIEEHLQRLTOINRMATP 92

## RESULT 15

ID Q95ZWL PRELIMINARY; PRT; 234 AA.  
 AC Q95ZWL;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Hypothetical protein F08F8.8.  
 GN ORFNames=F08F8.8;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Wilson R.;  
 RT "genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium";  
 RL Science 282:2012-2016(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Du Z.;  
 RT Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL EMBL; U28991; AAK6831.1; -  
 DR WormPep; F08F8.8; CE27935.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0006886; P:intracellular protein transport; IEA.  
 DR InterPro; IPR007705; V-SNARE.  
 DR Pfam; PF05008; V-SNARE; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 234 AA; 26776 MW; A3176FE9F31EA53 CRC64;

Query Match 35.1%; Score 61; DB 2; Length 234;  
 Best Local Similarity 42.9%; Pred. No. 4.1;  
 Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 8 FETMAIEIQLARLTGVNDKMAEYTN 35  
 DB 51 FXTVTEIEGLIEQLTINDDNDVAGA 78

Search completed: November 10, 2004, 12:26:55  
 Job time : 41.2948 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 ; Search time 9.8287 Seconds  
(without alignments)  
222.664 Million cell updates/sec

Title: US-10-092-750-8

Sequence: 1 AVOEDPVQREIHQDWMANREYIITSSIKIKAD 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/aa/5A.COMB.pap:\*  
2: /cgn2\_6/ptodata/1/aa/5B.COMB.pap:\*  
3: /cgn2\_6/ptodata/1/aa/6A.COMB.pap:\*  
4: /cgn2\_6/ptodata/1/aa/6B.COMB.pap:\*  
5: /cgn2\_6/ptodata/1/aa/PCTUS.COMB.pap:\*  
6: /cgn2\_6/ptodata/1/aa/backfile1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	162	95.9	75	4	US-09-489-847-206
2	162	95.9	75	4	US-09-621-976-7222
3	162	95.9	75	4	US-09-621-976-7223
4	162	95.9	75	4	US-09-621-976-7224
5	162	95.9	75	4	US-09-621-976-7225
6	162	95.9	75	4	US-09-621-976-7226
7	162	95.9	75	4	US-09-621-976-7227
8	53	31.4	351	4	US-09-809-665A-157
9	52	30.8	87	2	US-08-332-562A-92
10	50	29.6	227	4	US-09-328-352-6446
11	50	29.6	560	2	US-08-981-690-2
12	49	29.0	548	4	US-09-538-092-1058
13	49	29.0	648	4	US-09-107-532A-5531
14	48	28.4	128	4	US-09-107-532A-4132
15	48	28.4	255	4	US-09-540-236-3420
16	48	28.4	261	4	US-09-245-764-7
17	48	28.4	358	4	US-09-710-279-918
18	48	28.4	362	3	US-09-134-001C-5355
19	48	28.4	409	4	US-09-533-029-104
20	48	28.4	708	2	US-08-576-165-2
21	48	28.4	708	4	US-09-443-780C-16
22	48	28.4	708	4	US-09-079-723-110
23	48	28.4	708	4	US-09-079-723-116
24	48	28.4	2710	1	US-08-480-604A-6
25	48	28.4	2710	2	US-08-405-486A-6
26	48	28.4	2710	3	US-08-915-136-6
27	48	28.4	2710	3	US-08-957-310-6

28	48	28.4	2710	4	US-10-011-366-6	Sequence 6, Appli
29	48	28.4	2710	4	US-09-084-517-6	Sequence 6, Appli
30	47.5	28.1	535	4	US-09-543-681A-4593	Sequence 4593, Ap
31	47	27.8	199	3	US-09-078-317-2	Sequence 2, Appli
32	47	27.8	199	3	US-09-454-818-2	Sequence 2, Appli
33	47	27.8	574	4	US-08-248-786A-20132	Sequence 20132, A
34	47	27.8	582	1	US-08-261-086-2	Sequence 2, Appli
35	47	27.8	582	1	US-08-261-086-6	Sequence 6, Appli
36	47	27.8	582	1	US-08-261-086-8	Sequence 8, Appli
37	47	27.8	877	4	US-09-328-352-8162	Sequence 8162, Ap
38	47	27.8	1725	4	US-09-562-702A-20	Sequence 20, Appli
39	47	27.8	1725	4	US-09-561-818A-20	Sequence 20, Appli
40	47	27.8	1786	4	US-09-562-702A-18	Sequence 18, Appli
41	47	27.8	1786	4	US-08-561-818A-18	Sequence 18, Appli
42	46.5	27.5	148	4	US-09-489-039A-9258	Sequence 9258, Ap
43	46.5	27.5	225	4	US-09-543-681A-6942	Sequence 6942, Ap
44	46.5	27.5	543	4	US-09-535-008-63	Sequence 63, Appli
45	46.5	27.5	577	4	US-09-535-008-61	Sequence 61, Appli

# ALIGNMENTS

```

RESULT 1
US-09-489-847-206
; Sequence 206, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; EARLIER FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 206
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-206

Query Match          95.9%; Score 162; DB 4; Length 75;
Best Local Similarity 97.0%; Pred. No. 7.8e-17;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 AVOEDPVQREIHQDWMANREYIITSSIKIKAD 33  
 DB 2 AVOEDPVQREIHQDWMANREYIITSSIKIKAD 34

```

RESULT 2
US-09-621-976-7222
; Sequence 7222, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976

```

```

: RESULT 7
: US-09-621-976-7227
: Sequence 7227, Application US/09621976
: Patent No. 6639063
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J. B.
: APPLICANT: Jobert, S.
: APPLICANT: Giordano, J. Y.
: TITLE OF INVENTION: ESTs and Encoded Human Proteins
: FILE REFERENCE: GENSER, 054PR2
: CURRENT APPLICATION NUMBER: US/09/621,976
: CURRENT FILING DATE: 2000-07-21

```

NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO: 7227  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-621-976-7227

Query Match 95.9%; Score 162; DB 4; Length 75;  
Best Local Similarity 97.0%; Pred. No. 7.8e-17;  
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVQEDPVQREIHQDWMANREYIITSSIKKIAD 33  
DB 2 AGQEDPVQREIHQDWMANREYIITSSIKKIAD 34

RESULT 8  
US-09-809-665A-157  
Sequence 157, Application US/09809665A  
Patent No. 6790950  
GENERAL INFORMATION:  
APPLICANT: Lowery E., David, et al.  
TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions  
FILE REFERENCE: 28341/00435  
CURRENT APPLICATION NUMBER: US/09/809/665A  
CURRENT FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: 60/153,453  
PRIOR FILING DATE: 1999-09-10  
PRIOR APPLICATION NUMBER: 60/128,689  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 09/545,199  
PRIOR FILING DATE: 2000-04-06  
NUMBER OF SEQ ID NOS: 197  
SOFTWARE: Patentlin Ver. 2.0  
SEQ ID NO: 157  
LENGTH: 351  
TYPE: PRT  
ORGANISM: Actinobacillus pleuropneumoniae  
US-09-809-665A-157

Query Match 31.4%; Score 53; DB 4; Length 351;  
Best Local Similarity 53.3%; Pred. No. 8;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 6 PVQREIHQDWMANREY 20  
DB 117 PVKEINQDWMNKPY 131

RESULT 9  
US-08-332-562A-92  
Sequence 92, Application US/08332562A  
Patent No. 5985599  
GENERAL INFORMATION:  
APPLICANT: MCKENZIE, Ian F.C.  
APPLICANT: HOGARTH, Mark P.  
APPLICANT: HIBBS, Margaret L.  
APPLICANT: SCOTT, Bernadette M.  
APPLICANT: BONADONNA, Lisa  
APPLICANT: HOLETT, Mark D.  
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN  
NUMBER OF SEQUENCES: 136  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible

US-08-332-562A-92  
Sequence 92, Application US/08332562A  
Patent No. 5985599  
GENERAL INFORMATION:  
APPLICANT: MCKENZIE, Ian F.C.  
APPLICANT: HOGARTH, Mark P.  
APPLICANT: HIBBS, Margaret L.  
APPLICANT: SCOTT, Bernadette M.  
APPLICANT: BONADONNA, Lisa  
APPLICANT: HOLETT, Mark D.  
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN  
NUMBER OF SEQUENCES: 136  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentlin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/332,562A  
FILING DATE: 31-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/896,457  
FILING DATE: 27-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 54270/119/GRHA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 92:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-332-562A-92

Query Match 30.8%; Score 52; DB 2; Length 87;  
Best Local Similarity 53.3%; Pred. No. 2.2;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 AVQEDPVQREIHQDWM 15  
DB 72 SMSPDPVQLQIHNDW 86

RESULT 10  
US-09-328-352-6446  
Sequence 6446, Application US/09328352  
Patent No. 6562938  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO: 6446  
LENGTH: 227  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-6446

Query Match 29.6%; Score 50; DB 4; Length 227;  
Best Local Similarity 40.9%; Pred. No. 13;  
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 11 IHQDWMANREYIITSSIKKIA 32  
DB 62 IKXPMAGQGVENITPEIRRIA 83

RESULT 11  
US-08-981-690-2  
Sequence 2, Application US/08981690  
Patent No. 5976844  
GENERAL INFORMATION:  
APPLICANT: KAESLER, Bruno, Hermann, STAHMANN, Klaus-  
APPLICANT: Peter, SCHMIDT, Georg, BOEDDECKER, Theo;  
APPLICANT: SEUBLERGER, Harald  
TITLE OF INVENTION: PREPARATION OF RIBOFLAVIN USING  
TITLE OF INVENTION: MICROORGANISMS HAVING AN ALTERED  
TITLE OF INVENTION: ISOCITRATE LYASE ACTIVITY  
NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kell & Weinlauf  
STREET: 1101 Connecticut Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage  
COMPUTER: IBM AT-compatible, 60486 processor  
OPERATING SYSTEM: MS-DOS version 6.1  
SOFTWARE: WordPerfect version 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/981,690  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/590,454  
FILING DATE: 23-JUN-1996  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 560 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-981-690-2

Query Match 29.6%; Score 50; DB 2; Length 560;  
Best Local Similarity 39.4%; Pred. No. 39;  
Matches 13; Conservative 4; Mismatches 4; Indels 12; Gaps 2;

QY 8 QREI-----HQMAREYEITSSIKIA 32  
DB 505 QREMEGVVLKHQKAGAEYVD---SILKIA 533

RESULT 12  
US-09-538-092-1058  
Sequence 1058; Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Glot, Iolc  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuraParseqFormatter Version 0.9  
SEQ ID NO 1058  
LENGTH: 548  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: Polypeptide Accession Number P32780  
US-09-538-092-1058

Query Match 29.0%; Score 49; DB 4; Length 548;  
Best Local Similarity 31.6%; Pred. No. 53;  
Matches 12; Conservative 6; Mismatches 8; Indels 12; Gaps 1;

QY 2 VOEDPYQREIHOD-----WMAREYEITSS 27  
DB 118 LOEDPVLFLYKDLVVSQVISAEPWANLNVNATDSS 155

RESULT 13

US-09-107-532A-5531  
Sequence 5531; Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5531:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 648 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...648  
SEQUENCE DESCRIPTION: SEQ ID NO: 5531:  
US-09-107-532A-5531

Query Match 29.0%; Score 49; DB 4; Length 648;  
Best Local Similarity 32.0%; Pred. No. 65;  
Matches 8; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 3 QEDPYQREIHQMAREYEITSS 27  
DB 218 EEPVINEHYDGGRKSYVEYLMAN 242

RESULT 14  
US-09-107-532A-4132  
Sequence 4132; Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts

Fri Nov 12 14:56:03 2004

us-10-092-750-8.ra1

Page 5

Search completed: November 10, 2004, 12:32:23  
Job time : 10.8287 secs

```

; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...128
; SEQUENCE DESCRIPTION: SEQ ID NO: 4132:
;
US-09-107-532A-4132
;
Query Match 28.4%; Score 48; DB 4; Length 128;
Best Local Similarity 28.1%; Pred. No. 14;
Matches 9; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 2 VOEDPVQREIHODMANREVIETISIKKIAD 33
Db 49 IQEDIVFRSLQEDFMDVQNEHTLSAALSRLTD 80

RESULT 15
US-09-540-236-3420
; Sequence 3420: Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3420
; LENGTH: 255
; TYPE: PRT
; ORGANISM: M. catarrhalis
;
US-09-540-236-3420
;
Query Match 28.4%; Score 48; DB 4; Length 255;
Best Local Similarity 34.5%; Pred. No. 31;
Matches 10; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 5 DPVQREIHODMANREVIETISIKKIAD 33
Db 163 DEAKAQIKQTLAKQKATIELAMADAKGVAD 191
```

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 ; Search time 31.8287 Seconds  
(without alignments)  
366.225 Million cell updates/sec

Title: US-10-092-750-8

Perfect score: 169

Sequence: 1 AVQEDPVQREIHQDWANREYIHTTSIKKIAD 33

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35322586 residues 1566620

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	169	100.0	33 14 US-10-092-750-8	Sequence 8, Appl1
2	162	99.9	75 15 US-10-351-334-206	Sequence 206, Appl1
3	162	99.9	75 15 US-10-353-616-389	Sequence 389, App
4	162	99.9	110 14 US-10-092-750-243	Sequence 243, App
5	59	34.9	153 16 US-10-767-701-40435	Sequence 40435, A
6	59	34.9	727 17 US-10-425-115-339380	Sequence 339380,
7	55	32.5	84 17 US-10-425-115-334633	Sequence 334633,
8	55	32.5	84 17 US-10-425-115-334636	Sequence 334636,
9	55	32.5	85 15 US-10-424-599-157592	Sequence 157592,
10	55	32.5	86 16 US-10-437-963-116149	Sequence 116149,
11	55	32.5	131 15 US-10-425-114-50890	Sequence 50890, A
12	55	32.5	137 15 US-10-424-599-157593	Sequence 157593,
13	55	32.5	149 16 US-10-437-963-198829	Sequence 198829,

14	55	32.5	336 14 US-10-027-736A-65	Sequence 65, Appl1
15	55	32.5	357 14 US-10-027-736A-9	Sequence 9, Appl1
16	55	32.5	741 16 US-10-437-963-127661	Sequence 127661,
17	54	32.0	99 9 US-09-867-550-1368	Sequence 1368, Ap
18	54	32.0	261 14 US-10-094-749-2539	Sequence 2539, Ap
19	53	31.4	78 17 US-10-425-115-356318	Sequence 356318,
20	53	31.4	351 11 US-09-809-665A-157	Sequence 157, App
21	52	30.8	228 17 US-10-425-115-278588	Sequence 278588,
22	52	30.8	228 15 US-10-374-780A-1531	Sequence 1531, Ap
23	52	30.8	228 15 US-10-412-699B-1635	Sequence 1635, Ap
24	52	30.8	228 16 US-10-437-963-149962	Sequence 149962,
25	52	30.8	251 14 US-10-319-799-68	Sequence 68, Appl1
26	52	30.8	331 14 US-10-369-493-1104	Sequence 1104, Ap
27	52	30.8	341 15 US-10-382-122A-46828	Sequence 46828, A
28	52	30.8	404 15 US-10-369-493-21853	Sequence 21853, A
29	52	30.8	404 15 US-10-193-377-7	Sequence 7, Appl1
30	52	30.8	452 17 US-10-425-115-204655	Sequence 204655,
31	52	30.8	478 15 US-10-425-114-64779	Sequence 64779, A
32	51.5	30.5	113 16 US-10-437-963-195371	Sequence 195371,
33	51	30.2	300 14 US-10-374-780A-458	Sequence 458, App
34	51	30.2	300 15 US-10-374-780A-458	Sequence 46112, A
35	51	30.2	1302 15 US-10-282-122A-46112	Sequence 13591, A
36	50	29.6	397 9 US-09-815-242-13591	Sequence 254828,
37	50	29.6	688 15 US-10-424-599-254828	Sequence 22207, A
38	50	29.6	1442 14 US-10-369-493-22207	Sequence 245679,
39	49.5	29.3	266 17 US-10-425-115-245679	Sequence 245682,
40	49.5	29.3	266 17 US-10-425-115-245682	Sequence 53311, A
41	49.5	29.3	287 15 US-10-425-114-53311	Sequence 63394, A
42	49.5	29.3	287 15 US-10-425-114-53394	Sequence 57645, A
43	49	29.0	648 15 US-10-382-122A-57645	Sequence 17, Appl1
44	49	29.0	707 14 US-10-087-402-17	Sequence 1309, Ap
45	49	29.0	1514 16 US-10-408-765A-1309	

## ALIGNMENTS

RESULT 1  
US-10-092-750-8  
Sequence 8, Application US/10092750  
Publication No. US20030032157A1  
GENERAL INFORMATION:  
APPLICANT: Hammond, Philip W.  
APPLICANT: Alpin, Julia  
TITLE OF INVENTION: Polypeptides Interactive with BCL-X1  
FILE REFERENCE: 50036/050002  
CURRENT APPLICATION NUMBER: US/10/092,750  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: US 60/274,526  
PRIOR FILING DATE: 2001-03-08  
NUMBER OF SEQ ID NOS: 253  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 33  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-092-750-8.

Query Match 100.0%; Score 169; DB 14; Length 33;  
Best Local Similarity 100.0%; Pred. No. 1.4e-16;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVQEDPVQREIHQDWANREYIHTTSIKKIAD 33  
DB 1 AVQEDPVQREIHQDWANREYIHTTSIKKIAD 33

RESULT 2  
US-10-351-334-206  
Sequence 206, Application US/10351334  
Publication No. US20040034196A1  
GENERAL INFORMATION:

```

; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031p2
; CURRENT APPLICATION NUMBER: US/10/351,334
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 206
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-351-334-206

Query Match          95.9%; Score 162; DB 15; Length 75;
Best Local Similarity 97.0%; Pred. No. 3.5e-15;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVOEDPVOREIHODMANREYIEIITSSIKKIAD 33
Db 2 AGOEDPVOREIHODMANREYIEIITSSIKKIAD 34

RESULT 3
; US-10-363-616-389
; Sequence 389, Application US/10363616
; Publication No. US20040044181A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-113 (793)
; CURRENT APPLICATION NUMBER: US/10/363,616
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 09/654,935
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 490
; SEQ ID NO 389
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-363-616-389

Query Match          95.9%; Score 162; DB 15; Length 75;
Best Local Similarity 97.0%; Pred. No. 3.5e-15;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVOEDPVOREIHODMANREYIEIITSSIKKIAD 33
Db 2 AGOEDPVOREIHODMANREYIEIITSSIKKIAD 34

```

```

; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 243
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-092-750-243

Query Match          95.9%; Score 162; DB 14; Length 110;
Best Local Similarity 97.0%; Pred. No. 5.5e-15;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVOEDPVOREIHODMANREYIEIITSSIKKIAD 33
Db 6 AGOEDPVOREIHODMANREYIEIITSSIKKIAD 38

RESULT 5
; US-10-767-701-40435
; Sequence 40435, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40435
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C61471_1.pep
; US-10-767-701-40435

Query Match          34.9%; Score 59; DB 16; Length 153;
Best Local Similarity 47.6%; Pred. No. 2.8;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 5 DVOREIHODMANREYIEIIT 25
Db 30 DGVIEIHOHMKKEVVKVIT 50

RESULT 6
; US-10-425-115-339380
; Sequence 339380, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 339380
; LENGTH: 727
; TYPE: PRT

```

ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_72685C.1.pep  
US-10-425-115-334630

Query Match 34.8%; Score 59; DB 17; Length 727;  
Best Local Similarity 47.6%; Pred. No. 17;  
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 5 DPVREIHQDMAREYIEIIT 25  
Db 592 DGVEIEIHQDMAREYIEIIT 612

RESULT 7  
US-10-425-115-334633  
Sequence 334633, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 334633  
LENGTH: 84  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_68299C.1.pep  
US-10-425-115-334633

Query Match 32.5%; Score 55; DB 17; Length 84;  
Best Local Similarity 34.8%; Pred. No. 5.2;  
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy 11 IHQDMAREYIEIITSSIKKIAD 33  
Db 17 VQADWNEREFLISLNVRLFD 39

RESULT 8  
US-10-425-115-334636  
Sequence 334636, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 334636  
LENGTH: 84  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_68300C.1.pep  
US-10-425-115-334636

Query Match 32.5%; Score 55; DB 17; Length 84;  
Best Local Similarity 34.8%; Pred. No. 5.2;  
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy 11 IHQDMAREYIEIITSSIKKIAD 33  
Db 17 VQADWNEREFLISLNVRLFD 39

RESULT 9  
US-10-424-599-157592  
Sequence 157592, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 157592  
LENGTH: 85  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_113325C.1.pep  
US-10-424-599-157592

Query Match 32.5%; Score 55; DB 15; Length 85;  
Best Local Similarity 34.8%; Pred. No. 5.2;  
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy 11 IHQDMAREYIEIITSSIKKIAD 33  
Db 17 VQADWNEREFLISLNVRLFD 39

RESULT 10  
US-10-437-963-116149  
Sequence 116149, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbaruk, Brad  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 116149  
LENGTH: 86  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_13678C.1.pep  
US-10-437-963-116149

Query Match 32.5%; Score 55; DB 16; Length 86;  
Best Local Similarity 47.4%; Pred. No. 5.3;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 7 VQREIHQDMAREYIEIIT 25  
Db 8 VIREIHQDMAREYIEIIT 26

## RESULT 11

US-10-425-114-50890  
 ; Sequence 50890, Application US/10425114  
 ; Publication No. US2004003468A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 50890  
 ; LENGTH: 131  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700076923\_Flt.pep  
 US-10-425-114-50890

Query Match 32.5%; Score 55; DB 15; Length 131;  
 Best Local Similarity 34.8%; Pred. No. 8.6;  
 Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 11 IHODWANEYIEITSSIKKIAD 33  
 : ||||| : : : : :  
 Db 30 VQADWNERFISINVRRLFD 52

## RESULT 12

US-10-424-599-157593  
 ; Sequence 157593, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J  
 ; APPLICANT: Kovalic, David K  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 157593  
 ; LENGTH: 137  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(137)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_113326C.1.pep  
 US-10-424-599-157593

Query Match 32.5%; Score 55; DB 15; Length 137;  
 Best Local Similarity 34.8%; Pred. No. 9.1;  
 Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 11 IHODWANEYIEITSSIKKIAD 33  
 : ||||| : : : : :  
 Db 35 VQADWNERFISINVRRLFD 57

RESULT 13  
 US-10-437-963-198829  
 ; Sequence 198829, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 198829  
 ; LENGTH: 149  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_94450C.1.pep  
 US-10-437-963-198829

Query Match 32.5%; Score 55; DB 16; Length 149;  
 Best Local Similarity 34.8%; Pred. No. 10;  
 Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 11 IHODWANEYIEITSSIKKIAD 33  
 : ||||| : : : : :  
 Db 19 VQADWNERFISINVRRLFD 41

## RESULT 14

US-10-027-736A-65  
 ; Sequence 65, Application US/10027736A  
 ; Publication No. US20030190614A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Presta, Leonard G.  
 ; APPLICANT: Nemenuk, Angela K.  
 ; TITLE OF INVENTION: NON-HUMAN PRIMATE FC RECEPTORS AND METHODS OF USE  
 ; FILE REFERENCE: 11669.92US01  
 ; CURRENT APPLICATION NUMBER: US/10/027,736A  
 ; CURRENT FILING DATE: 2002-08-23  
 ; NUMBER OF SEQ ID NOS: 72  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 65  
 ; LENGTH: 336  
 ; TYPE: PRT  
 ; ORGANISM: Cynomolgus  
 ; FEATURE:  
 ; NAME/KEY: MISC\_FEATURE  
 ; LOCATION: (1)..(336)  
 ; OTHER INFORMATION: FcgammaRII alpha-chain  
 US-10-027-736A-65

Query Match 32.5%; Score 55; DB 14; Length 336;  
 Best Local Similarity 61.5%; Pred. No. 25;  
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEDPVOREIHQDW 15  
 : ||||| : : : : :  
 Db 71 RSDPIQLRIHRDW 83

## RESULT 15

US-10-027-736A-9  
 ; Sequence 9, Application US/10027736A  
 ; Publication No. US20030190614A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Presta, Leonard G.  
 ; APPLICANT: Nemenuk, Angela K.  
 ; TITLE OF INVENTION: NON-HUMAN PRIMATE FC RECEPTORS AND METHODS OF USE

Fri Nov 12 14:56:03 2004

us-10-092-750-8.rapb

Page 5

```

1  FILE REFERENCE: 11669.92US01
2  CURRENT APPLICATION NUMBER: US/10/027,736A
3  CURRENT FILING DATE: 2002-08-23
4  NUMBER OF SEQ ID NOS: 72
5  SOFTWARE: PatentIn version 3.1
6  SEQ ID NO 9
7  LENGTH: 357
8  TYPE: PRT
9  ORGANISM: Cynomolgus
10 FEATURE:
11 NAME/KEY: MISC FEATURE
12 LOCATION: (1)..(357)
13 OTHER INFORMATION: FcgammaRI <chain
14 US-10-027-736A-9

```

Query Match	32.5%	Score 55;	DB 14;	Length 357;
Best Local Similarity	61.5%	Pred. No. 27;		
Matches	8;	Conservative	3;	Mismatches 2;
				Indels 0;
				Gaps 0;
Qx	3	GGDPVQRIRHDM	15	
Db	92	KSDPLQLEIRHDM	104	

```
Search completed: November 10, 2004, 16:35:44
Job time : 32.8287 secs
```

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:36:51; Search time 6.6713 Seconds  
(without alignments)  
475.942 Million cell updates/sec

Title: US-10-092-750-8

Perfect score: 169  
Sequence: 1 AVQEDPVQREIHQDMANREYEITTSIKKIAD 33

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56.5	33.4	173	1 B45628	ferritin heavy cha
2	56	33.1	900	2 T50773	translactin initia
3	55	32.5	900	2 T47732	probable translati
4	54	32.0	290	2 T37151	probable DNA-bindi
5	54	32.0	381	2 T48623	hypothetical prote
6	52	30.8	242	2 G83747	dethiodictin synth
7	52	30.8	331	2 F69025	ATP phosphoribosyl
8	52	30.8	351	2 JCA4752	ferrochelatase (EC
9	52	30.8	404	2 A45480	Fc gamma (IgG) rec
10	52	30.8	414	2 AG1614	B. subtilis Yx10 p
11	52	30.8	414	2 A11251	B. subtilis Yx10 p
12	51.5	30.5	222	2 T38461	ubiquitin carboxyl
13	51.5	30.5	997	2 S26426	type IV site-speci
14	51	30.2	55	2 E97235	uncharacterized pr
15	51	30.2	61	2 A84615	hypothetical prote
16	50	29.6	147	2 S74460	hypothetical prote
17	50	29.6	250	2 D69215	conserved hypothet
18	50	29.6	250	2 S30584	hypothetical prote
19	50	29.6	493	2 F90746	hypothetical prote
20	50	29.6	493	2 B85597	probable sulfatase
21	50	29.6	495	2 AC0607	probable sulfatase
22	50	29.6	873	2 E90581	hypothetical prote
23	50	29.6	1036	2 D84741	probable cellulose
24	50	29.6	1442	2 S57160	sulfite reductase
25	49.5	29.3	280	2 T55577	Pc gamma (IgG) rec
26	49.5	29.3	521	2 T37252	probable matrix me
27	49.5	29.3	555	2 F70372	dihydroxyacid deny
28	49	29.0	253	2 S74480	hypothetical prote
29	49	29.0	266	2 S44954	Imbg protein - Str

30	49	29.0	548	2 S27958	transcription fact
31	49	29.0	614	1 S10032	urease (EC 3.5.1.5
32	48.5	28.7	409	2 T31662	hypothetical prote
33	48.5	28.7	741	2 T12711	NADH2 dehydrogenas
34	48.5	28.7	741	2 T13764	NADH2 dehydrogenas
35	48.5	28.7	757	2 C70034	conserved hypothet
36	48.5	28.7	3027	2 J01917	polyprotein - pars
37	48	28.4	216	2 D98240	probable hydrolase
38	48	28.4	229	2 AH3045	hydrolase (importe
39	48	28.4	344	2 A41357	Fc gamma (IgG) rec
40	48	28.4	374	1 A39878	Fc gamma (IgG) rec
41	48	28.4	656	2 T47373	hypothetical prote
42	48	28.4	708	2 A56163	peptide transport
43	48	28.4	718	2 T51488	hypothetical prote
44	48	28.4	992	2 E75093	translactin initia
45	48	28.4	1151	2 T04657	hypothetical prote

## ALIGNMENTS

## RESULT 1

B45628 ferritin heavy chain 1 - fluke (Schistosoma mansoni)

C/Species: Schistosoma mansoni

C/Date: 22-Apr-1993 #sequence\_revision 02-Jun-1994 #text\_change 09-Jul-2004

C/Accession: B45628

R/Dietzel, J.; Hirzmann, J.; Preis, D.; Symmons, P.; Kunz, W.

Mol. Biochem. Parasitol. 50, 245-254, 1992

A/Title: Ferritins of Schistosoma mansoni: sequence comparison and expression in female

A/Reference number: A45628; MUID:92158004; PMID:1741011

A/Accession: B45628

A/Molecule type: mRNA

A/Residues: 1-173 <DIE>

A/Cross-references: UNIPROT:P25319; GB:M4538; NID:G160985; PIDN:AAA29880.1; PID:G16096

A/Note: Sequence extracted from NCBI database (NCBIN:82751, NCBIPI:82752)

C/Suprafamily: ferritin

C/Keywords: iron; iron storage; metalloprotein

F,23,57,58,60,61,103/Binding site: iron (Glu, Glu, Glu, Glu, His, Glu) #status Predicted

Query Match 33.4%; Score 56.5; DB 1; Length 173;  
Best Local Similarity 46.2%; Pred. No. 1.8;  
Matches 12; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

Cy 9 REIH-QDMANREYEITTSIKKIAD 33

Db 121 RDTHFCDPFINNEYLEIQVSMKKLSD 146

## RESULT 2

T50773 translactin initiation factor homolog p105 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000. #text\_change 09-Jul-2004

C/Accession: T50773

A/Title: The Arabidopsis homologue of an eIF3 complex subunit associates with the COP9

A/Reference number: 224448; MUID:9849901; PMID:9849901

A/Accession: T50773

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-900 <KAR>

A/Cross-references: UNIPROT:Q49160; EMBL:AF040102; PIDN:AA03464.1

Query Match 33.1%; Score 56; DB 2; Length 900;  
Best Local Similarity 37.9%; Pred. No. 14;  
Matches 11; Conservative 8; Mismatches 8; Indels 2; Gaps 1;

Cy 2 VQEDPVQREIHQDMANREYEITTSIKK 30

Db 230 NKQDP--KEITWQVNNKKEIVAAAREK 256

## RESULT 3

T47732

Probable translation initiation factor eIF-3 chain 8 F18021.110 [imported] - Arabidopsis

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C:Accession: T47732

R:Benne, V.; Wurmbach, E.; Dizonck, H.; Amseorge, W.; Mewes, H.W.; Rudd, S.; Lemcke, K.;

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24474

A:Accession: T47732

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-900 &lt;BN&gt;

A:Cross-references: UNIPROT:Q49160; EMBL:AF18021

A:Experimental source: cultivar Columbia; BAC clone F18021

C:Genetics:

A:Map position: 3

A:Introns: 8/3; 593/3

A&gt;Note: F18021.110

## Query Match

Best Local Similarity 32.5%; Score 55; DB 2; Length 900;

Matches 11; Conservative 8; Mismatches 2; Indels 2; Gaps 1;

QY 2 VQDPVQREIHODWVANNREYIEITSSIKK 30

DB 230 MNKDP--KEITWDVWKKRKEIVAAKGGK 256

## RESULT 4

T37151

Probable DNA-binding protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T37151

R:Harries, D.; Bentley, S.D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21622

A:Accession: T37151

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-290 &lt;HAR&gt;

A:Cross-references: UNIPROT:Q9S105; EMBL:AL109972; PDB:1CAB53276.1; GSPDB:GN00070; SCOE

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SCJ9A.15C

## Query Match

Best Local Similarity 32.0%; Score 54; DB 2; Length 290;

Matches 9; Conservative 8; Mismatches 7; Indels 2; Gaps 1;

QY 5 DPVQREIHODWVANNREYIEITSSIKK 28

DB 157 DPVQREIHODWVANNREYIEITSSIKK 182

## RESULT 5

T46623

Hypothetical protein F18022.290 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C:Accession: T46623

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24493

A:Accession: T46623

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-381 &lt;BEV&gt;

A:Cross-references: UNIPROT:Q9LV79; EMBL:DB

A:Experimental source: cultivar Columbia; BAC clone F18022

C:Genetics:

A:Map position: 5

A:Introns: 50/3; 100/3; 125/3; 144/3; 197/2; 241/2; 275/3; 309/1

A&gt;Note: F18022.290

C:Superfamily: Arabidopsis thaliana hypothetical protein F9D16.200

Query Match

Best Local Similarity 32.0%; Score 54; DB 2; Length 381;

Matches 8; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 EDPVQREIHODWVANNREYIEITSS 27

DB 342 ESPVILKPEHMKRGQELSISSSS 365

## RESULT 6

G83747

dethiobiotin synthetase BH0783 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004

C:Accession: G83747

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: G83747

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-242 &lt;STO&gt;

A:Cross-references: UNIPROT:Q3KER6; GB:AP001509; GB:BA000004; NID:G10173176; PIDN:BA04

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH0783

C:Superfamily: dethiobiotin synthetase

## Query Match

Best Local Similarity 30.8%; Score 52; DB 2; Length 242;

Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 4 EDPVQREIHODWVANNREYIEITSS 26

DB 219 DDKLRRLHYDMDKRRFKELIHS 241

## RESULT 7

F69025

ATP phosphoribosyltransferase related protein - Methanobacterium thermoautotrophicum (s

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C:Accession: F69025

R:Smith, D.R.; Doucette-Stamm, L.A.; Delouche, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qi, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, F.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: F69025

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-331 &lt;MTH&gt;

A:Cross-references: UNIPROT:Q26222; GB:AE000801; GB:AE000666; NID:G2621145; PIDN:AA846

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH119

A:Start codon: GTG

## Query Match

Best Local Similarity 30.8%; Score 52; DB 2; Length 331;

Matches 16; Conservative 2; Mismatches 13; Indels 8; Gaps 2;

QY 3 QEDPVQ--RE-----IHODWVANNREYIEITSSIKKIAD 33

DB 62 QSAPEVLENNEMDLALIGEDWVRESINSNGSIKKIGD 100



## RESULT 8

UC4752  
 ferriochelataase (EC 4.99.1.1) - Rhodobacter capsulatus  
 C:Species: Rhodobacter capsulatus  
 C:Date: 10-May-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004  
 C:Accession: J04752  
 R:Kamazireva, E.; Biel, A.J.  
 Gene 170, 149-150, 1996  
 A:Title: Nucleotide sequence of the Rhodobacter capsulatus hemH gene.  
 A:Reference number: J04752; MUID:96200875; PMID:8621079  
 A:Accession: J04752  
 A:Molecule type: DNA  
 A:Residues: 1-351 <KAN>  
 A:Cross-references: UNIPROT:Q59735; GB:U034391; NID:g1002892; PIDN:AAA88884.1; PID:g10028  
 C:Genetics:  
 A:Gene: hemH  
 A:Function:  
 A:Description: catalyzes the insertion of iron into protoporphyrin to produce heme  
 C:Superfamily: ferriochelataase  
 C:Keywords: heme biosynthesis; iron; lyase  
 F:214-233/Region: ferriochelataase motif

Query Match 30.8%; Score 52; DB 2; Length 351;  
 Best Local Similarity 30.0%; Pred. No. 17;  
 Matches 9; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

OY 2 VOEDPVQREIHQDMANREYIEITTSIKKI 31  
 DB 175 VKRQPAARTVPEYFARPSYIEALASSVERV 204

## RESULT 9

A46480  
 Fc gamma (IgG) receptor high affinity - mouse  
 N:Alternate names: high affinity IgG receptor  
 C:Species: Mus musculus (house mouse)  
 C:Date: 18-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C:Accession: A46480; A43511  
 R:Osman, N.; Kozak, C.A.; McKenzie, I.F.; Hogarth, P.M.  
 J. Immunol. 148, 1570-1575, 1992  
 A:Title: Structure and mapping of the gene encoding mouse high affinity Fc gamma RI and  
 A:Reference number: A46480; MUID:92165399; PMID:1531670  
 A:Accession: A46480  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-404 <OSM>  
 A:Cross-references: UNIPROT:P26151  
 A:Note: sequence extracted from NCBI backbone (NCBIN:85205, NCBIN:85208, NCBIN:85212, NC  
 R:Seares, D.W.; Osman, N.; Tate, B.; McKenzie, I.F.C.; Hogarth, P.M.  
 J. Immunol. 144, 371-378, 1990  
 A:Title: Molecular cloning and expression of the mouse high affinity Fc receptor for IgG  
 A:Reference number: A43511; MUID:90111035; PMID:2136886  
 A:Accession: A43511  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-404 <SEA>  
 A:Cross-references: GB:M31314; NID:g200752; PIDN:AAA40056.1; PID:g200753  
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology  
 C:Keywords: immunoglobulin receptor; transmembrane protein  
 F:127-179/Domain: immunoglobulin homology <IMM>

Query Match 30.8%; Score 52; DB 2; Length 404;  
 Best Local Similarity 53.3%; Pred. No. 20;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 AVOEDPVQREIHQDM 15  
 DB 100 SWSDPVQQLINDW 114

## RESULT 10

## AG1614

B. subtilis Yx10 protein homolog lin1456 [imported] - Listeria innocua (strain C1p1126  
 C:Species: Listeria innocua  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AG1614  
 R:Glaser, P.; Franguel, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feith, F  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Matounam, A.; B  
 ok, C.; Schluecker, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlan  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AG1614  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-414 <GLA>  
 A:Cross-references: UNIPROT:Q92B78; GB:AL592022; PIDN:CAC96687.1; PID:g16413929; GSPDB:  
 A:Experimental source: strain C1p11262  
 C:Genetics:  
 A:Gene: lin1456

Query Match 30.8%; Score 52; DB 2; Length 414;  
 Best Local Similarity 52.4%; Pred. No. 20;  
 Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 11 IHQDMANREYIEITTSIKKI 31  
 DB 11 IFQDMANSAYSIMTTALPI 31

## RESULT 11

A11251  
 B. subtilis Yx10 protein homolog lmo1417 [imported] - Listeria monocytogenes (strain EC  
 C:Species: Listeria monocytogenes  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: A11251  
 R:Glaser, P.; Franguel, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feith, F  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Matounam, A.; B  
 ok, C.; Schluecker, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlan  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: A11251  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-414 <GLA>  
 A:Cross-references: UNIPROT:Q8Y779; GB:NC\_003210; PIDN:CAC99495.1; PID:g16410846; GSPDB  
 A:Experimental source: strain EGD-e  
 C:Genetics:  
 A:Gene: lmo1417

Query Match 30.8%; Score 52; DB 2; Length 414;  
 Best Local Similarity 52.4%; Pred. No. 20;  
 Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 11 IHQDMANREYIEITTSIKKI 31  
 DB 11 IFQDMANSAYSIMTTALPI 31

## RESULT 12

T38461  
 ubiquitin carboxyl-terminal hydrolase isozyme 13 - fission yeast (Schizosaccharomyces I  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T38461  
 R:Harris, D.; McDonald, S.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, February 1996  
 A:Reference number: Z21794  
 A:Accession: T38461

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-222 <HAR>  
A:Cross-references: UNIPROT:Q10171; EMBL:Z69368; PIDN:CAA93292.1; GSPDB:GN00066; SPDB:SH  
A:Experimental source: strain 97zh-; cosmid c27F1  
C:Genetics:  
A:Gene: SPDB:SPAC27F1.03c  
A:Map position: 1  
C:Superfamily: human ubiquitin thioesterase

Query Match 30.5%; Score 51.5; DB 2; Length 222;  
Best Local Similarity 43.3%; Pred. No. 12;  
Matches 13; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

OY 3 QEDPVOR-RIHODMANREYIEITSSIKKI 31  
DB 180 QEGPVQSEIESDLNNAEVLVSVIKYIQSI 209

## RESULT 13

S26426  
Type IV site-specific deoxyribonuclease (EC 3.1.21.-) Eco57I - Escherichia coli  
N/Alternate names: endonuclease Eco57I; type IV restriction modification system Eco 57I  
N/Contributing site-specific DNA-methyltransferase (adenine-specific) (EC 2.1.1.72) R.Eco57  
C/Species: Escherichia coli  
C/Date: 22-Nov-1993 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
A/Accession: S26426; S26679  
R/Tanulaitis, A.; Valisvita, R.; Timinskas, A.; Kimauskas, S.; Butkus, V.  
Submitted to the EMBL Data Library, July 1991  
A/Description: Cloning and sequence determination of the genes coding for Eco57I type IV  
A/Reference number: S26424  
A/Accession: S26426  
A:Molecule type: DNA  
A:Residues: 1-997 <JAN>  
A/Cross-references: UNIPROT:P25239; EMBL:X61122; NID:G41318; PIDN:CAA43434.1; PID:G41321  
R/Tanulaitis, A.; Valisvita, R.; Timinskas, A.; Kimauskas, S.; Butkus, V.  
Nucleic Acids Res. 20, 6051-6056, 1992  
A/Title: Cloning and sequence analysis of the genes coding for Eco57I type IV restriction  
A/Reference number: S26677; NID:9309595; PMID:1334261  
A/Accession: S26679  
A:Molecule type: DNA  
A:Residues: 388-400; 520-534; 556-595 <JAN>  
A/Cross-references: EMBL:X61122  
A/Note: the authors state that the endonuclease polypeptide is bifunctional and has both  
C/Function: <<END>>  
A/Description: type IV site-specific deoxyribonuclease Eco57I; restriction endonuclease  
A/Function: <<MER>>  
A/Description: site-specific DNA-methyltransferase (adenine-specific) R.Eco57I; restrict  
A/Note: the R.Eco57I endonuclease contains DNA-methyltransferase activity; there is an  
C/Keywords: endonuclease; hydrolase; methyltransferase; multifunctional enzyme; restrict

Query Match 30.5%; Score 51.5; DB 2; Length 997;  
Best Local Similarity 35.5%; Pred. No. 65;  
Matches 11; Conservative 9; Mismatches 10; Indels 1; Gaps 1;

OY 4 EDPVOREI-HODMANR-EYIEITSSIKKIAD 33  
DB 923 DDPKQKEIHDTISSKQYINKLVSOTQKAD 953

## RESULT 14

E97235  
Uncharacterized protein, probable yncB. subtilis homolog [imported] - Clostridium acet  
C/Species: Clostridium acetobutylicum  
C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C/Accession: E97235  
R/Molling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
A/Reference number: A96900; NID:21359325; PMID:21359325  
A/Accession: E97235  
A>Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-55 <KUR>  
A/Cross-references: UNIPROT:Q97FLO; GB:AE001437; PIDN:AAK80672.1; PID:G15025761; GSPDB  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC2726

Query Match 30.2%; Score 51; DB 2; Length 55;  
Best Local Similarity 44.8%; Pred. No. 2; 9;  
Matches 13; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

OY 3 QEDPVOR-RIHODMANREYIEITSSIK 29  
DB 18 KEDGLTEEEKKQDITLRREYIEITKGNV 46

## RESULT 15

A84615  
Hypochemical protein Atg22640 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: A84615  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Xoo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; NID:20083487; PMID:1061197  
A/Accession: A84615  
A/Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-61 <STD>  
A/Cross-references: UNIPROT:Q9ZQ49; GB:AE002093; NID:G4314361; PIDN:AAID15572.1; GSPDB:C  
A:Gene: Atg22640  
A:Map position: 2

Query Match 30.2%; Score 51; DB 2; Length 61;  
Best Local Similarity 33.3%; Pred. No. 3; 2;  
Matches 7; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 11 IHODMANREYIEITSSIKKI 31  
DB 17 VQADWENREFTSHISLVRL 37

Search completed: November 10, 2004, 12:29:09  
Job time : 7.6713 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:27:01 ; Search time 36.1065 Seconds  
(without alignments)  
525.871 Million cell updates/sec

Title: US-10-092-750-8  
Perfect score: 169  
Sequence: 1 AVQEDPVQREIHQDMANREYIEITTSIKIAD 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: uniprot\_spprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	162	95.9	Q8WUW1	Q8WUW1 homo sapien
2	162	95.9	AAN60161	AAN60161 homo sapi
3	162	95.9	AAN60162	AAN60162 homo sapi
4	162	95.9	AAN60167	AAN60167 homo sapi
5	162	95.9	Q9P082	Q9P082 homo sapien
6	161	95.3	Q61086	Q61086 brachydactio
7	159	94.1	Q91VR8	Q91VR8 mus musculi
8	159	94.1	Q6P7G6	Q6P7G6 xenopus lae
9	159	94.1	AAN61677	AAN61677 xenopus l
10	121	71.6	Q8MLQ0	Q8MLQ0 drosophila
11	64	37.9	Q70G88	Q70G88 anopheles g
12	60	35.5	Q8E852	Q8E852 leptospira
13	60	35.5	Q72MT6	Q72MT6 leptospira
14	60	35.5	Q8E852	Q8E852 leptospira
15	59	34.9	Q90276	Q90276 chimpanzee
16	59	34.9	Q9FYT6	Q9FYT6 zea mays (m
17	56.5	33.4	FRH1 SCHMA	FRH1 SCHMA
18	55	32.5	Q8RW98	Q8RW98 zea mays (m
19	55	32.5	Q8RW98	Q8RW98 zea mays (m
20	55	32.5	Q8SPW5	Q8SPW5 macaca fasc
21	55	32.5	Q61YR3	Q61YR3 oryza sativ
22	55	32.5	BA03815	BA03815 oryza sat
23	55	32.5	IF38 ARATH	IF38 ARATH
24	55	32.5	9001	9001 Arabidopsis
25	55	32.5	Q6H8X5	Q6H8X5 bacillus th
26	54	32.0	Q8C953	Q8C953 mus musculi
27	54	32.0	Q8BH10	Q8BH10 mus muscu
28	54	32.0	Q8BH10	Q8BH10 mus muscu
29	54	32.0	Q96SN7	Q96SN7 homo sapien
30	54	32.0	AAN69270	AAN69270 homo sapi
31	54	32.0	Q6NZ16	Q6NZ16 xenopus lae

32	54	32.0	257	2	AAN66119
33	54	32.0	290	2	Q9S105
34	54	32.0	306	2	Q84W23
35	54	32.0	349	2	Q9M270
36	54	32.0	381	2	Q9LY79
37	53.5	31.7	269	2	Q93T01
38	53.5	31.7	312	2	Q9G9N5
39	53.5	31.7	321	2	Q9G9N5
40	53.5	31.7	323	2	Q9G9N5
41	53	31.4	364	2	Q51848
42	53	31.4	463	2	Q7PVK3
43	53	31.4	637	2	Q7LYC6
44	53	31.4	637	2	Q7LYW2
45	53	31.4	637	2	Q9HGZ8

## ALIGNMENTS

RESULT 1  
ID Q8WUW1 PRELIMINARY; PRT; 75 AA.  
AC Q8WUW1;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)  
DE Chromosome 3 open reading frame 10 (Brlt-like protein).  
GN Name=C3orf10;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
EX MEDLINE=22386257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
RA Ditschenko L., Marziani K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stempleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Breenstein M.J., Ueda T.B., Tomihata S., Carninci P., Franke C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouford G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
RA Krzyzanski M.I., Skalska U., Small D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
EX Strausberg R.;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain, and Lung;  
EX Xu Q., Duan R., Huo Y., Fan B., Zhang K., Wu D.;  
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
EX Strausberg R.;  
RA Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RL EMBL; BC019303; AAN61903.1; -  
DR EMBL; AY148219; AAN60161.1; -

DR EMBL: AY148220; AAN60162.1; -  
 DR EMBL: BC001067; AAH01067.1; -  
 SQ SEQUENCE 75 AA; 8745 MW; 637B5461A2D96587 CRC64;

Query Match 95.9%; Score 162; DB 2; Length 75;  
 Best Local Similarity 97.0%; Pred. No. 2,3e-14;  
 Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVGEDPVQREIHQDMANREYIEITSSIKKIAD 33  
 DB 2 AGGEDPVQREIHQDMANREYIEITSSIKKIAD 34

RESULT 2

AAH01067 PRELIMINARY; PRT; 75 AA.  
 AC AAN60161;  
 DT 02-MAR-2004 (TEMBLrel. 27, Created)  
 DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)  
 DE Btk1-like protein.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primata; Catarrhini; Homiidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Xu Q., Duan R., Huo Y., Fan B., Zhang K., Wu D.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY148219; AAN60161.1; -  
 SQ SEQUENCE 75 AA; 8745 MW; 637B5461A2D96587 CRC64;

Query Match 95.9%; Score 162; DB 2; Length 75;  
 Best Local Similarity 97.0%; Pred. No. 2,3e-14;  
 Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVGEDPVQREIHQDMANREYIEITSSIKKIAD 33  
 DB 2 AGGEDPVQREIHQDMANREYIEITSSIKKIAD 34

RESULT 3

AAH01067 PRELIMINARY; PRT; 75 AA.  
 AC AAN60162;  
 DT 02-MAR-2004 (TEMBLrel. 27, Created)  
 DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)  
 DE Btk1-like protein.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primata; Catarrhini; Homiidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Petal brain;  
 RA Xu Q., Duan R., Huo Y., Fan B., Zhang K., Wu D.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY148220; AAN60162.1; -  
 SQ SEQUENCE 75 AA; 8745 MW; 637B5461A2D96587 CRC64;

Query Match 95.9%; Score 162; DB 2; Length 75;  
 Best Local Similarity 97.0%; Pred. No. 2,3e-14;  
 Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVGEDPVQREIHQDMANREYIEITSSIKKIAD 33  
 DB 2 AGGEDPVQREIHQDMANREYIEITSSIKKIAD 34

RESULT 4  
 AAH01067

ID AAH01067 PRELIMINARY; PRT; 75 AA.  
 AC AAH01067;  
 DT 10-MAY-2004 (TEMBLrel. 27, Created)  
 DT 10-MAY-2004 (TEMBLrel. 27, Last sequence update)  
 DE Hypothetical protein.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=22368257; Pubmed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan B., Moore T., Wax S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Kravynski M.I., Skalska U., Smailus D.E., Schurch A., Schein J.E.,  
 RA Jones S., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Strausberg R.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC001067; AAH01067.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 75 AA; 8745 MW; 637B5461A2D96587 CRC64;

Query Match 95.9%; Score 162; DB 2; Length 75;  
 Best Local Similarity 97.0%; Pred. No. 2,3e-14;  
 Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVGEDPVQREIHQDMANREYIEITSSIKKIAD 33  
 DB 2 AGGEDPVQREIHQDMANREYIEITSSIKKIAD 34

RESULT 5

Q9P082 PRELIMINARY; PRT; 110 AA.  
 AC Q9P082;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DE HSPC300 (Fragment).  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,  
 RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF161418; AAF28978.1; -  
 FT NON TER 1  
 SQ SEQUENCE 110 AA; 12376 MW; 96254639D78E33F1 CRC64;

Query Match 95.9%; Score 162; DB 2; Length 110;  
 Best Local Similarity 97.0%; Pred. No. 3,5e-14;  
 Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AVQEDPVQREIHQDANREYIIITSSIKKIAD 33  
 DB 6 AGQEDPVQREIHQDANREYIIITSSIKKIAD 38

RESULT 6  
 ID 061086 PRELIMINARY; PRT; 75 AA.  
 AC 061086;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DE Zgc:86903.  
 GN Name=zgc:86903;  
 OS Brachydanio rerio (zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stalton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Utsin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX Strausberg R.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC071524; AAH71524.1; -;  
 SQ SEQUENCE 75 AA; 8731 MW; 637B5476A9B4C457 CRC64;

Query Match 95.3%; Score 161; DB 2; Length 75;  
 Best Local Similarity 93.9%; Pred. No. 3,1e-14;  
 Matches 31; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AVQEDPVQREIHQDANREYIIITSSIKKIAD 33  
 DB 2 AGQEDPVQREIHQDANREYIIITSSIKKIAD 34

RESULT 7  
 ID 091VR8 PRELIMINARY; PRT; 75 AA.  
 AC 091VR8;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE RIKEN CDNA 6720456B07 (Mus musculus 12 days embryo male wolffian duct

DE includes surrounding region cDNA, RIKEN full-length enriched library,  
 DE clone:6720456B07 product:hypothetical protein, full insert  
 DE sequence).  
 GN Name=6720456B07Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Mammary tumor. Brca1-/fl;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stalton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Utsin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Mammary tumor. Brca1-/fl;  
 RX Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DE includes surrounding region cDNA, RIKEN full-length enriched library,  
 DE clone:6720456B07 product:hypothetical protein, full insert  
 DE sequence).  
 GN Name=6720456B07Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Mammary tumor. Brca1-/fl;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stalton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Utsin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Mammary tumor. Brca1-/fl;  
 RX Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujitake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multiplexed capillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6J; TISSUE=Mojillian duct includes surrounding region;  
 RA Adachi U., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Haragaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imocani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Shiraki T.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Tanaka T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.,  
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC010582; AAH10582.1; -  
 DR EMBL: AK020128; BAC25616.1; -  
 DR MGD: MG1:1915406; 6720456B07Rik.  
 DR Hypothetical protein.  
 KW SEQUENCE  
 SQ 75 AA; 8761 MW; 637B43D779A36587 CRC64;  
 Query Match 94.1%; Score 159; DB 2; Length 75;  
 Best Local Similarity 93.9%; Pred. No. 5,9e-14;  
 Matches 31; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AVOEDPVQREIHQDMANREYIEITSSIKKIAD 33  
 DB 2 AVOEDPVQREIHQDMANREYIEITSSIKKIAD 34  
 RESULT 8  
 QY 06P7G6 PRELIMINARY; PRT; 75 AA.  
 ID 06P7G6;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE MGC68812 protein.  
 GN Name:MGC68812;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Scheffer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.C., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalios D.E., Scherch A., Schein J.E.,

RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT Initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Klein S., Strausberg R.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC016777; AAH61677.1; -  
 DR EMBL: BC016777; AAH61677.1; -  
 DR MGD: MG1:1915406; 6720456B07Rik.  
 KW SEQUENCE  
 SQ 75 AA; 8747 MW; 872F5422FC0C5B6 CRC64;  
 Query Match 94.1%; Score 159; DB 2; Length 75;  
 Best Local Similarity 90.9%; Pred. No. 5,9e-14;  
 Matches 30; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AVOEDPVQREIHQDMANREYIEITSSIKKIAD 33  
 DB 2 AVOEDPVQREIHQDMANREYIEITSSIKKIAD 34  
 RESULT 9  
 QY AAH61677 PRELIMINARY; PRT; 75 AA.  
 ID AAH61677;  
 DT 14-APR-2004 (TrEMBLrel. 27, Created)  
 DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 14-APR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE MGC68812 protein.  
 GN Name:MGC68812;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Scheffer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.C., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalios D.E., Scherch A., Schein J.E.,

RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Klein S., Strausberg R.;  
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC061677; AAH61677.1; -  
SQ SEQUENCE 75 AA; 8747 MW; 872F5422FCF0C556 CRC64;  
  
Query Match 94.4%; Score 159; DB 2; Length 75;  
Best Local Similarity 90.9%; Pred. No. 5.9e-14;  
Matches 30; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 AVOEDPQREIHODMANREYIEITTSIKKIAD 33  
ID 08MLQ0 PRELIMINARY; PRT; 76 AA.  
AC 08MLQ0;  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE CG30173-PA.  
GN ORFNames=CG30173;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidae; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,  
RA Abril J.F., Aghayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktiroglu U., Beasley E.M.,  
RA Beeson K.Y., Bencs P.V., Bernan B.P., Bhargava D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck H., Brokstein P., Broctler P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Chertiy J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Durbin K.J., Evanske C.C., Ferraz C., Ferrier S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pauleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson R., Skupski M.P., Smith T.,  
RA Spieler E., Spredling A.C., Stapleton M., Strong R., Sun E.,  
RA Svrtk R., Tector A., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Dugan M., Champe M., Lavery T., Muzny D.M., Nelson C.R.,  
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
RA Pauleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svrtk R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
RT melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svrtk R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
RT a genomic perspective.";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktiroglu U., Betman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review.";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RG FLYBASE;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE003462; AAH68291.1; -  
DR FLYBASE: FBgn0061198; STPL.  
SQ SEQUENCE 76 AA; 8812 MW; 59947DEC8E05437A CRC64;  
  
Query Match 71.6%; Score 121; DB 2; Length 76;  
Best Local Similarity 69.0%; Pred. No. 9.5e-09;  
Matches 20; Conservative 7; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 DPVQREIHODMANREYIEITTSIKKIAD 33  
ID 07Q388 PRELIMINARY; PRT; 7667 AA.  
AC 07Q388;  
DT 01-MAR-2004 (TREMBlrel. 26, Created)  
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE AC014535 (Fragment).  
GN Name=egCG49728; ORFNames=ENSHNG00000017478;  
OS Anopheles gambiae str. PST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anophelinae.  
OX NCBI\_Taxid=180454;  
RN [1]  
RP SEQUENCE FROM N.A.

```

RC STRAIN=PEST;
RA Arophaeas Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
DR EMBL; AAB01006339; EAA5893.1; -
DR InterPro; IPR001715; Calpainin-like.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00435; Spectrin; 18.
DR PROSITE; PSS0021; CH; 2.
FT NON_TER
FT NON_TER
SQ SEQUENCE 7667 AA; 883480 MW; F423B58BC8HDE577 CRC64;

Query Match
Best Local Similarity 37.9%; Score 64; DB 2; Length 7667;
Matches 13; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 2 VOEDPVQREIHQDMANREYIEITTSIKKIAD 33
Db 6086 VRKDDQGETEDYRLRELEKLGQSLKKVID 6117

RESULT 12
PRELIMINARY; PRT; 1396 AA.
ID QSEZ52;
AC QSEZ52;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Adenylate cyclase (EC 4.6.1.1).
GN Name=CyA16; OrderedLocNames=LA4008;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OC NCBI_TaxID=173;
RX NCBI_TaxID=173;
RP SEQUENCE FROM N.A.
RA STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-O., Jia J., Tu Y.-F.,
RA Zhang Y.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Yao Z.-Y., Shen Y., Qiang B.-Q., Qian Z., Wang S.-Y., Ma W.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "unique physiological and pathogenic features of Leptospira
RT Interrogans revealed by whole-genome sequencing.";
RL Nature 422:688-693(2003).
DR EMBL; AEO11556; AAN51206.1; -
DR GO; GO:0004016; F:adenylate cyclase activity; IEA.
DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001054; G_cyclase.
DR InterPro; IPR001440; TPR_1like.
DR InterPro; IPR008941; TPR_1like.
DR SMART; SM00044; CYCC; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PSS0125; GUANYLATE_CYCLASES_2; 2.
DR PROSITE; PSS0005; TPR; 1.
DR PROSITE; PSS0293; TPR_REGION; 1.
DR Complete proteome; Lyase.
SQ SEQUENCE 1396 AA; 157621 MW; E1A9FA323704B2E0 CRC64;

Query Match
Best Local Similarity 35.5%; Score 60; DB 2; Length 1396;
Matches 13; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 AVQEDPVQREIHQDMANREYIEITTSIKKIAD 33
Db 1396 VRKDDQGETEDYRLRELEKLGQSLKKVID 6117

```

```

Db 579 SVEEDPLTREIDRKQNERLEFIITSLQERAN 611

RESULT 13
PRELIMINARY; PRT; 1398 AA.
ID Q72MI6;
AC Q72MI6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Adenylate/guanylate cyclase.
GN OrderedLocNames=L1c13201;
OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar
OS Copenhagen).
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OC NCBI_TaxID=44275;
RP SEQUENCE FROM N.A.
RA STRAIN=Fiocruz LI-130;
RX PubMed=15028702;
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,
RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,
RA Coutinho L.L., Degraive W.M., Dellegostin O.A., El-Dorry H.,
RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gambierini M., Gigliotti E.A.,
RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
RA Jerônimo S.M.B., Junqueira-de-Azevedo J.L.M., Kimura E.T.,
RA Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
RA de Oliveira R.C., Pereira G.G., Reis M.S., Schriener A.,
RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
RT "Comparative genomics of two Leptospira interrogans serovars reveals
RT novel insights into physiology and pathogenesis.";
RL J. Bacteriol. 186:2164-2172(2004).
DR EMBL; AEO17300; AAS71746.1; -
DR InterPro; IPR001054; G_cyclase.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR_1like.
DR SMART; SM00028; TPR; 4.
DR PROSITE; PSS0125; GUANYLATE_CYCLASES_2; 2.
DR PROSITE; PSS0005; TPR; 1.
DR PROSITE; PSS0293; TPR_REGION; 1.
SQ Complete proteome.
SQ SEQUENCE 1398 AA; 157793 MW; F16CBB403D58B114 CRC64;

Query Match
Best Local Similarity 35.5%; Score 60; DB 2; Length 1398;
Matches 13; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 AVQEDPVQREIHQDMANREYIEITTSIKKIAD 33
Db 579 SVEEDPLTREIDRKQNERLEFIITSLQERAN 611

RESULT 14
PRELIMINARY; PRT; 1398 AA.
ID AAS71746;
AC AAS71746;
DT 24-MAR-2004 (TrEMBLrel. 27, Created)
DT 24-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 11-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Adenylate/guanylate cyclase.
GN L1c13201.
OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar
OS Copenhagen).
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OC NCBI_TaxID=44275;
RP SEQUENCE FROM N.A.
RA STRAIN=Fiocruz LI-130;
RX PubMed=15028702;
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,

```



Job time : 39.1065 secs

RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrier H.,  
RA Coutinho L.L., Degraeve W.M., Dellagostin O.A., El-Dorri H.,  
RA Ferro E.S., Ferro M.T., Furlan L.R., Gamberini M., Gigliotti E.A.,  
RA Gees-Neco A., Goldman G.H., Goldman W.H.S., Harakava R.,  
RA Uertonio S.M.B., Junqueira-de-Azevedo I.L.N., Kinura E.T.,  
RA Kurihara E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,  
RA de Oliveira R.C., Pereira G.G., Reis M.S., Schrieler A.,  
RA Siqueira L.E.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,  
RA Camargo L.R.A., Kitajima J.P., Seubald J.C., Van Sluys M.A.,  
RT "Comparative genomics of two *Leptospira interrogans* serovars reveals  
RT novel insights into physiology and pathogenesis.",  
RL J. Bacteriol. 186:2164-2172(2004).  
DR ENBL173001, MAS17746.1, --, F16CB6403D58B114 CRC64,  
QO SEQUENCE 1396 AA; 157793 MW; 157793 MW; 157793 MW;

Query Match	35.5%;	Score 60;	DB 2;	Length 1398;
Best Local Similarity	39.4%;	Pred. No. 48;		
Matches 13; Conservative	8;	Mismatches 12;	Indels 0;	Gaps 0;

QY 1 AVQEDPVQGEIHQDMANREYIEITSSIKIAD 33  
: : ||| : ||| : : :  
Db 579 SVEEDPLTREIDRKQKNERLFEIITSSLGERAN 611

## RESULT 15

ID	030276	PRELIMINARY;	PRT;	101 AA.
AC	030276;			
DT	01-NOV-1998	(TrEMBLrel. 08, Created)		
DT	01-NOV-1998	(TrEMBLrel. 08, Last sequence update)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)		
DE	Tat protein.			
GN	Name=tat;			
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_taxid=11723;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SiVhoest;			
RX	MEDLINE=99098990; PubMed=9882304;			
RA	Hirsch V.M., Campbell B.J., Balles E., Goeken R., Brown C.,			
RT	Elkins W.R., Axthelm M., Murphy-cord M., Sharp P.M.,			
RT	"Characterization of a novel simian immunodeficiency virus (SIV) from			
RT	l'hoest monkeys (Cercopithecus l'hoestii): implications for the origins			
RT	of SIVmd and other primate lentiviruses.";			
RL	J. Virol. 73:1036-1045 (1999).			
CC	-1- FUNCTION: Transcriptional regulator that acts by binding to the			
CC	trans-activating responsive sequence (TAR) RNA element and			
CC	activates transcription initiation and/or elongation from the LTR			
CC	promoter (by similarity).			
DR	EMBL; AF075269; AAD1251.1; -.			
DR	HSSP; P12506; ITBC.			
DR	GO: GO:0042025; C:host cell nucleus; IEA.			
DR	GO: GO:0005634; C:nucleus; IEA.			
DR	GO: GO:0003723; F:RNA binding; IEA.			
DR	GO: GO:0003700; F:transcription factor activity; IEA.			
DR	GO: GO:0006355; P:regulation of transcription, DNA-depends; IEA.			
DR	Interpro: IPR001831; IV_Tat.			
DR	Pfam: PF00539; Tat; 1.			
DR	PRINTS: PRO0055; HIVTATDOMAIN.			
KM	Activator; Nuclear protein; RNA-binding; Transcription regulation.			
QO	SEQUENCE 101 AA; 12015 KM; 3F19G11AA;D9953C C9C64;			

Query Match	34.9%	Score 59;	DB 2;	Length 101;
Best Local Similarity	37.0%	Pred. No. 3.9;		
Matches 10;	Conservative 7;	Mismatches 10;	Indels 0;	Gaps 0;

Qy 3 QEDPVOREITHQDMANREYIEITSSIK 29  
 6 QEQPIQKQHQLMKETYNEAUTKALQ 32

*This Page Blank (uspto)*

Fri Nov 12 14:56:04 2004

us-10-092-750-9.ra1

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 ; Search time 9.8287 Seconds  
(without alignments)  
222.664 Million cell updates/sec

Title: US-10-092-750-9

Perfect score: 162  
Sequence: 1 ATRQALNEISARHSGIQLEIRSLRDLHDIFFTL 33

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/aa/PCITUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	88.3	298	1	US-08-393-985-10 Sequence 10, Appl
2	103	63.6	161	3	US-08-493-071-17 Sequence 17, Appl
3	103	63.6	161	3	US-08-493-071-20 Sequence 20, Appl
4	103	63.6	187	3	US-08-493-071-16 Sequence 16, Appl
5	103	63.6	187	3	US-08-493-071-19 Sequence 19, Appl
6	103	63.6	236	3	US-08-493-071-15 Sequence 15, Appl
7	103	63.6	236	3	US-08-493-071-18 Sequence 18, Appl
8	103	63.6	277	1	US-08-690-457-5 Sequence 5, Appl
9	103	63.6	277	2	US-08-628-187-5 Sequence 5, Appl
10	103	63.6	279	1	US-08-493-071-3 Sequence 11, Appl
11	103	63.6	279	2	US-08-690-457-11 Sequence 11, Appl
12	103	63.6	279	2	US-08-628-187-11 Sequence 11, Appl
13	103	63.6	279	3	US-08-493-071-6 Sequence 6, Appl
14	103	63.6	287	1	US-08-690-457-4 Sequence 4, Appl
15	103	63.6	287	2	US-08-628-187-4 Sequence 4, Appl
16	103	63.6	287	3	US-08-493-071-2 Sequence 2, Appl
17	103	63.6	288	1	US-08-690-457-3 Sequence 3, Appl
18	103	63.6	288	1	US-08-690-457-10 Sequence 10, Appl
19	103	63.6	288	2	US-08-628-187-3 Sequence 3, Appl
20	103	63.6	288	2	US-08-628-187-10 Sequence 10, Appl
21	103	63.6	288	3	US-08-493-071-1 Sequence 1, Appl
22	103	63.6	288	3	US-08-493-071-5 Sequence 5, Appl
23	103	63.6	289	1	US-08-690-457-9 Sequence 9, Appl
24	103	63.6	289	2	US-08-628-187-9 Sequence 9, Appl
25	103	63.6	289	3	US-08-493-071-4 Sequence 4, Appl
26	103	63.6	290	1	US-08-393-985-6 Sequence 6, Appl
27	100	61.7	288	1	US-08-356-397-2 Sequence 2, Appl

28	100	61.7	288	1	US-08-393-985-4 Sequence 4, Appl
29	95	58.6	263	1	US-08-393-985-23 Sequence 23, Appl
30	95	58.6	285	1	US-08-393-985-2 Sequence 2, Appl
31	95	58.6	288	1	US-08-337-602-4 Sequence 4, Appl
32	95	58.6	288	3	US-08-558-135-4 Sequence 4, Appl
33	95	58.6	288	3	US-08-819-286-3 Sequence 3, Appl
34	89	54.9	279	3	US-08-819-286-5 Sequence 5, Appl
35	89	54.9	289	3	US-08-393-985-8 Sequence 8, Appl
36	89	54.9	289	3	US-08-819-286-6 Sequence 6, Appl
37	87	53.7	28	3	US-08-819-286-29 Sequence 29, Appl
38	87	53.7	31	1	US-08-393-985-21 Sequence 21, Appl
39	87	53.7	50	1	US-08-393-985-19 Sequence 19, Appl
40	87	53.7	76	1	US-08-393-985-13 Sequence 13, Appl
41	73	45.1	24	3	US-08-819-286-28 Sequence 28, Appl
42	64	39.5	20	3	US-08-819-286-6 Sequence 6, Appl
43	64	39.5	20	3	US-08-819-286-27 Sequence 27, Appl
44	61	37.7	290	1	US-08-411-706-2 Sequence 2, Appl
45	58	35.8	295	1	US-08-411-706-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-08-393-985-10  
Sequence 10, Application US/08393985  
Patent No. 5693476  
GENERAL INFORMATION:  
APPLICANT: Scheller, Richard H.  
TITLE OF INVENTION: Methods and Compositions for Modulation  
of Vesicular Release  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESS: Dellinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,985  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Shultz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 8600-0152  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 298 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-393-985-10

Query Match 88.3%; Score 143; DB 1; Length 298;  
Best Local Similarity 90.6%; Pred. No. 2.3e-12;  
Matches 29; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

CY 2 TROALNEISARHSGIQLEIRSLRDLHDIFFTL 33  
DB 196 TROALNEISARHSGIQLEIRSLRDLHDIFFTL 227

RESULT 2

```

US-08-493-071-17
; Sequence 17, Application US/08493071
; Patent No. 6127149
; GENERAL INFORMATION:
; APPLICANT: Hirai, Yohel
; APPLICANT: Koshida, Shogo
; APPLICANT: Oka, Yumiko
; TITLE OF INVENTION: MODIFIED EPIMORPHIN
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
; STREET: 99 CANAL CENTER PLAZA, SUITE 300
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/493,071
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Price, Robert L.
; REGISTRATION NUMBER: 22,685
; REFERENCE/DOCKET NUMBER: 715-107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-684-1111
; TELEFAX: 703-684-1124
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-493-071-17

Query Match      63.6%; Score 103; DB 3; Length 161;
Best Local Similarity 69.0%; Pred. No. 5e-07;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDI 30
DB      84 TROALNEISRHDIKMLTSIRELHEMF 112

RESULT 3
US-08-493-071-20
; Sequence 20, Application US/08493071
; Patent No. 6127149
; GENERAL INFORMATION:
; APPLICANT: Hirai, Yohel
; APPLICANT: Koshida, Shogo
; APPLICANT: Oka, Yumiko
; TITLE OF INVENTION: MODIFIED EPIMORPHIN
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
; STREET: 99 CANAL CENTER PLAZA, SUITE 300
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/08/493,071
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Price, Robert L.
; REGISTRATION NUMBER: 22,685
; REFERENCE/DOCKET NUMBER: 715-107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-684-1111
; TELEFAX: 703-684-1124
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-493-071-20

Query Match      63.6%; Score 103; DB 3; Length 161;
Best Local Similarity 69.0%; Pred. No. 5e-07;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDI 30
DB      84 TROALNEISRHDIKMLTSIRELHEMF 112

RESULT 4
US-08-493-071-16
; Sequence 16, Application US/08493071
; Patent No. 6127149
; GENERAL INFORMATION:
; APPLICANT: Hirai, Yohel
; APPLICANT: Koshida, Shogo
; APPLICANT: Oka, Yumiko
; TITLE OF INVENTION: MODIFIED EPIMORPHIN
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
; STREET: 99 CANAL CENTER PLAZA, SUITE 300
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/493,071
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Price, Robert L.
; REGISTRATION NUMBER: 22,685
; REFERENCE/DOCKET NUMBER: 715-107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-684-1111
; TELEFAX: 703-684-1124
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-493-071-16

Query Match      63.6%; Score 103; DB 3; Length 187;
Best Local Similarity 69.0%; Pred. No. 6e-07;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

Qy 2 TROALNEISARHSIGIOLETSIRELHDIF 30  
 Db 110 TROALNEISARHSIGIOLETSIRELHEMF 138

RESULT 5

US-08-493-071-19  
 ; Sequence 19, Application US/08493071  
 ; Patent No. 6127149  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hirai, Yohei  
 ; APPLICANT: Koshida, Shogo  
 ; APPLICANT: Oka, Yumiko  
 ; TITLE OF INVENTION: MODIFIED EPIMORPHIN  
 ; NUMBER OF SEQUENCES: 30  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER  
 ; STREET: 99 CANAL CENTER PLAZA, SUITE 300  
 ; CITY: ALEXANDRIA  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22314  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/493,071  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Price, Robert L.  
 ; REGISTRATION NUMBER: 22,685  
 ; REFERENCE/DOCKET NUMBER: 715-107  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-684-1111  
 ; TELEFAX: 703-684-1124  
 ; INFORMATION FOR SEQ ID NO: 19:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 187 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-493-071-19

Query Match 63.6%; Score 103; DB 3; Length 187;  
 Best Local Similarity 69.0%; Pred No. 6e-07; 5; Indels 0; Gaps 0;  
 Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSIGIOLETSIRELHDIF 30  
 Db 110 TROALNEISARHSIGIOLETSIRELHEMF 138

RESULT 6

US-08-493-071-15  
 ; Sequence 15, Application US/08493071  
 ; Patent No. 6127149  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hirai, Yohei  
 ; APPLICANT: Koshida, Shogo  
 ; APPLICANT: Oka, Yumiko  
 ; TITLE OF INVENTION: MODIFIED EPIMORPHIN  
 ; NUMBER OF SEQUENCES: 30  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER  
 ; STREET: 99 CANAL CENTER PLAZA, SUITE 300  
 ; CITY: ALEXANDRIA  
 ; STATE: VA  
 ; COUNTRY: USA

ZIP: 22314  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/493,071  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Price, Robert L.  
 ; REGISTRATION NUMBER: 22,685  
 ; REFERENCE/DOCKET NUMBER: 715-107  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-684-1111  
 ; TELEFAX: 703-684-1124  
 ; INFORMATION FOR SEQ ID NO: 15:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 236 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-493-071-15

Query Match 63.6%; Score 103; DB 3; Length 236;  
 Best Local Similarity 69.0%; Pred. No. 7.9e-07; 5; Indels 0; Gaps 0;  
 Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSIGIOLETSIRELHDIF 30  
 Db 159 TROALNEISARHSIGIOLETSIRELHEMF 187

RESULT 7

US-08-493-071-18  
 ; Sequence 18, Application US/08493071  
 ; Patent No. 6127149  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hirai, Yohei  
 ; APPLICANT: Koshida, Shogo  
 ; APPLICANT: Oka, Yumiko  
 ; TITLE OF INVENTION: MODIFIED EPIMORPHIN  
 ; NUMBER OF SEQUENCES: 30  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER  
 ; STREET: 99 CANAL CENTER PLAZA, SUITE 300  
 ; CITY: ALEXANDRIA  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22314  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/493,071  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Price, Robert L.  
 ; REGISTRATION NUMBER: 22,685  
 ; REFERENCE/DOCKET NUMBER: 715-107  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-684-1111  
 ; TELEFAX: 703-684-1124  
 ; INFORMATION FOR SEQ ID NO: 18:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 236 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single



```

APPLICATION NUMBER: US/08/493,071
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Price, Robert L.
REGISTRATION NUMBER: 22,685
REFERENCE/DOCKET NUMBER: 715-107
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-684-1111
TELEFAX: 703-684-1124
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-493-071-3

Query Match      63.6%; Score 103; DB 3; Length 277;
Best Local Similarity 69.0%; Pred. No. 9.5e-07;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      2 TROALNEISARHSGIOQLERSIRELHDI 30
DB      187 TROALNEISRHKDIKMLETISIRELHEMF 215

```

```

RESULT 11
US-08-690-457-11
Sequence 11, Application US/08690457
Patent No. 5726298
GENERAL INFORMATION:
APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.
TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE
TITLE OF INVENTION: DESIGNATED AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANTIBOD
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA,
STREET: 1, Taya-cho
CITY: Yokohama-shi
COUNTRY: Japan
ZIP: 244
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,457
FILING DATE: 16-AUG-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/078,309
FILING DATE: June 15, 1993
APPLICATION NUMBER: 294856/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 294857/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 122906/1992
FILING DATE: April 17, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 135692/1992
FILING DATE: April 30, 1992
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 279
TYPE: amino acid
TOPOLOGY: linear

```

```

MOLECULE TYPE: peptide
US-08-690-457-11

Query Match      63.6%; Score 103; DB 1; Length 279;
Best Local Similarity 69.0%; Pred. No. 9.6e-07;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      2 TROALNEISARHSGIOQLERSIRELHDI 30
DB      188 TROALNEISRHKDIKMLETISIRELHEMF 216

```

```

RESULT 12
US-08-628-187-11
Sequence 11, Application US/08628187
Patent No. 5837239
GENERAL INFORMATION:
APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.
TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE DESIGNATED
TITLE OF INVENTION: AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANTIBODIES THERE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA, Japan
STREET: 1, Taya-cho
CITY: Yokohama-shi
COUNTRY: Japan
ZIP: 244
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628,187
FILING DATE: April 5, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 294856/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 294857/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 122906/1992
FILING DATE: April 17, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 135692/1992
FILING DATE: April 30, 1992
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 279
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-628-187-11

Query Match      63.6%; Score 103; DB 2; Length 279;
Best Local Similarity 69.0%; Pred. No. 9.6e-07;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

RESULT 13
US-08-493-071-6
Sequence 6, Application US/08493071
Patent No. 6127149
GENERAL INFORMATION:
APPLICANT: Hirai, Yohel
APPLICANT: Koshida, Shogo
APPLICANT: Oka, Yumiko

```

TITLE OF INVENTION: MODIFIED EPIMORPHIN  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER  
STREET: 99 CANAL CENTER PLAZA, SUITE 300  
CITY: ALEXANDRIA  
STATE: VA  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/493,071  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Price, Robert L.  
REGISTRATION NUMBER: 22,685  
REFERENCE/DOCKET NUMBER: 715-107  
TELEPHONE: 703-684-1111  
TELEFAX: 703-684-1124  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 279 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-493-071-6

Query Match 63.6%; Score 103; DB 3; Length 279;  
Best Local Similarity 69.0%; Pred. No. 9.9e-07;  
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 TROLNEISARHSIGIOLESIREDHIF 30  
DB 188 TROLNEISRHDKIMKETSIRELHMF 216

RESULT 14  
US-08-690-457-4  
Sequence 4, Application US/08690457  
Patent No. 5726298  
GENERAL INFORMATION:  
APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.  
TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE  
TITLE OF INVENTION: DESIGNATED AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANTIBOD  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA,  
ADDRESS: Japan  
STREET: 1, Taya-cho  
CITY: Yokohama-shi  
COUNTRY: Japan  
ZIP: 244  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/690,457  
FILING DATE: 16-AUG-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/078,309  
FILING DATE: June 15, 1993  
APPLICATION NUMBER: 294856/1991

FILING DATE: October 16, 1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 294857/1991  
FILING DATE: October 16, 1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 122906/1992  
FILING DATE: April 17, 1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 135692/1992  
FILING DATE: April 30, 1992  
SEQUENCE CHARACTERISTICS:  
LENGTH: 287  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-628-187-4

Query Match 63.6%; Score 103; DB 1; Length 287;  
Best Local Similarity 69.0%; Pred. No. 9.9e-07;  
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 TROLNEISARHSIGIOLESIREDHIF 30  
DB 187 TROLNEISRHDKIMKETSIRELHMF 215

RESULT 15  
US-08-628-187-4  
Sequence 4, Application US/08628187  
Patent No. 5837239  
GENERAL INFORMATION:  
APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.  
TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE DESIGNATED  
TITLE OF INVENTION: AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANTIBODIES THERE  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA, Japan  
STREET: 1, Taya-cho  
CITY: Yokohama-shi  
COUNTRY: Japan  
ZIP: 244  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/628,187  
FILING DATE: April 5, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 294856/1991  
FILING DATE: October 16, 1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 294857/1991  
FILING DATE: October 16, 1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 122906/1992  
FILING DATE: April 17, 1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 135692/1992  
FILING DATE: April 30, 1992  
SEQUENCE CHARACTERISTICS:  
LENGTH: 287  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-628-187-4

Query Match 63.6%; Score 103; DB 2; Length 287;  
Best Local Similarity 69.0%; Pred. No. 9.9e-07;



Fri Nov 12 14:56:04 2004

us-10-092-750-9.rat

Page 7

Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 TROALNEISARHSGIOQLERSIRELHDIP 30  
Db 187 TROALNEIESRHRDKIMKLTSTIRELHEMF 215

Search completed: November 10, 2004, 12:32:23  
Job time : 9.8287 secs

THIS PAGE BLANK (USPIO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 ; Search time 31.8287 Seconds  
(without alignments)  
366.225 Million cell updates/sec

Title: US-10-092-750-9  
Perfect score: 162  
Sequence: 1 ATRQALNEISARHSIGQLERSIRELHDIFTL 33

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications\_AA.\*  
1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/1/pubpaa/PCR\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/1/pubpaa/PCRUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	162	100.0	33	US-10-092-750-9	Sequence 9, Appl1
2	150	92.6	295	US-10-104-047-2704	Sequence 2704, Ap
3	150	92.6	380	US-09-925-300-1492	Sequence 1492, Ap
4	103	63.6	272	US-10-364-141A-11	Sequence 11, Appl
5	103	63.6	287	US-10-364-141A-10	Sequence 10, Appl
6	103	63.6	288	US-10-364-141A-9	Sequence 9, Appl1
7	103	63.6	299	US-10-364-141A-12	Sequence 12, Appl
8	100	61.7	288	US-09-942-024-22	Sequence 22, Appl
9	100	61.7	288	US-09-942-098-22	Sequence 22, Appl
10	100	61.7	288	US-10-261-161-17	Sequence 17, Appl
11	95	58.6	259	US-10-408-765A-807	Sequence 807, Appl
12	95	58.6	288	US-09-942-024-21	Sequence 21, Appl
13	95	58.6	288	US-09-942-098-21	Sequence 21, Appl

14	95	58.6	288	US-10-261-161-16	Sequence 16, Appl
15	94	58.0	288	US-09-942-024-23	Sequence 23, Appl
16	94	58.0	288	US-09-942-098-23	Sequence 23, Appl
17	94	58.0	288	US-10-261-161-18	Sequence 18, Appl
18	88	54.3	288	US-10-466-113-1	Sequence 1, Appl1
19	88	54.3	291	US-09-942-024-25	Sequence 25, Appl
20	88	54.3	291	US-09-942-098-25	Sequence 25, Appl
21	88	54.3	291	US-10-261-161-20	Sequence 20, Appl
22	87	53.7	288	US-09-942-024-26	Sequence 26, Appl
23	87	53.7	288	US-09-942-098-26	Sequence 26, Appl
24	87	53.7	288	US-10-261-161-21	Sequence 21, Appl
25	84	51.9	291	US-09-942-024-24	Sequence 24, Appl
26	84	51.9	291	US-09-942-098-24	Sequence 24, Appl
27	84	51.9	291	US-10-261-161-19	Sequence 19, Appl
28	78	48.1	287	US-10-085-188-6	Sequence 6, Appl1
29	76	46.9	285	US-09-898-570-41	Sequence 41, Appl
30	76	46.9	285	US-09-898-570-42	Sequence 42, Appl
31	76	46.9	285	US-09-839-446-41	Sequence 41, Appl
32	76	46.9	285	US-09-839-446-42	Sequence 42, Appl
33	76	46.9	287	US-10-085-188-2	Sequence 2, Appl1
34	75	46.3	339	US-10-425-114-52515	Sequence 52515, A
35	72	44.4	305	US-10-739-930-64671	Sequence 64671, Ap
36	70	43.2	201	US-10-767-701-52121	Sequence 52121, A
37	70	43.2	314	US-10-425-115-355980	Sequence 355980, A
38	69	42.6	154	US-10-424-599-252190	Sequence 252190, A
39	69	42.6	330	US-10-437-963-113212	Sequence 113212, A
40	68	42.0	117	US-10-424-599-230736	Sequence 230736, A
41	67	41.4	308	US-10-437-963-170819	Sequence 170819, A
42	65	40.1	123	US-10-437-963-122175	Sequence 122175, A
43	65	40.1	311	US-10-437-963-130142	Sequence 130142, A
44	64	39.5	126	US-10-425-115-355977	Sequence 355977, A
45	64	39.5	294	US-09-898-570-28	Sequence 28, Appl

#### ALIGNMENTS

RESULT 1  
US-10-092-750-9  
; Sequence 9, Application US/10092750  
; Publication No. US2003032157A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; APPLICANT: Alpin, Julia  
; APPLICANT: Wright, Martin C.  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1  
; FILE REFERENCE: 50036/050002  
; CURRENT FILING DATE: US/10/092, 750  
; PRIOR APPLICATION NUMBER: 2002-03-07  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-750-9  
Query Match 100.0%; Score 162; DB 14; Length 33;  
Best Local Similarity 100.0%; Pred. No. 1.5e-15;  
Matches 33; Conservative 0; Mismatches 0; Indels 0;  
CQ 1 ATRQALNEISARHSIGQLERSIRELHDIFTL 33  
DB 1 ATRQALNEISARHSIGQLERSIRELHDIFTL 33  
RESULT 2  
US-10-104-047-2704  
; Sequence 2704, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:

```

; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392a1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2704
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2704

Query Match          92.6%; Score 150; DB 14; Length 295;
Best Local Similarity 96.9%; Pred. No. 9.4e-13;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDIFTL 33
DB      194 TROALNEISARHSEIOQLERSIRELHDIFTL 225

RESULT 3
US-09-925-300-1492
; Sequence 1492, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1492
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1492

Query Match          92.6%; Score 150; DB 9; Length 380;
Best Local Similarity 96.9%; Pred. No. 1.3e-12;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDIFTL 33
DB      279 TROALNEISARHSEIOQLERSIRELHDIFTL 310

RESULT 4
US-10-364-141A-11
; Sequence 11, Application US/10364141A
; Publication No. US20040091483A1
; GENERAL INFORMATION:
; APPLICANT: The Cleveland Clinic Foundation
; APPLICANT: Thomas, Weimbs
; TITLE OF INVENTION: SNARES AND METHODS OF CONTROLLING CYTOKINESIS
; FILE REFERENCE: 126233.1201
; CURRENT APPLICATION NUMBER: US/10/364,141A
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/355323
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 11
; LENGTH: 272
```

```

; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-364-141A-11

Query Match          63.6%; Score 103; DB 15; Length 272;
Best Local Similarity 69.0%; Pred. No. 3.5e-06;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDIF 30
DB      187 TROALNEISARHDKIMLETISRLEHMF 215

RESULT 5
US-10-364-141A-10
; Sequence 10, Application US/10364141A
; Publication No. US20040091483A1
; GENERAL INFORMATION:
; APPLICANT: The Cleveland Clinic Foundation
; APPLICANT: Thomas, Weimbs
; TITLE OF INVENTION: SNARES AND METHODS OF CONTROLLING CYTOKINESIS
; FILE REFERENCE: 126233.1201
; CURRENT APPLICATION NUMBER: US/10/364,141A
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/355323
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 10
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-364-141A-10

Query Match          63.6%; Score 103; DB 15; Length 287;
Best Local Similarity 69.0%; Pred. No. 3.7e-06;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDIF 30
DB      187 TROALNEISARHDKIMLETISRLEHMF 215

RESULT 6
US-10-364-141A-9
; Sequence 9, Application US/10364141A
; Publication No. US20040091483A1
; GENERAL INFORMATION:
; APPLICANT: The Cleveland Clinic Foundation
; APPLICANT: Thomas, Weimbs
; TITLE OF INVENTION: SNARES AND METHODS OF CONTROLLING CYTOKINESIS
; FILE REFERENCE: 126233.1201
; CURRENT APPLICATION NUMBER: US/10/364,141A
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/355323
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 9
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-364-141A-9

Query Match          63.6%; Score 103; DB 15; Length 288;
Best Local Similarity 69.0%; Pred. No. 3.7e-06;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDIF 30
DB      187 TROALNEISARHDKIMLETISRLEHMF 215
```



Best Local Similarity 65.5%; Pred. No. 4.4e-05;  
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
Qy 2 TROALNEISARHSGIOQLERSIRELHDF 30  
Db 159 SKQALSEIETRHSEIITKLENSIRELHDMF 187

RESULT 12  
US-09-942-024-21  
; Sequence 21, Application US/09942024  
; Publication No. US20030143650A1  
; GENERAL INFORMATION:  
; APPLICANT: Steward, Lance E.  
; APPLICANT: Fernandez-Salas, Ester  
; APPLICANT: Aoki, Kei Roger  
; TITLE OF INVENTION: Fret Protease Assays For Botulinum  
; TITLE OF INVENTION: Serotype A/E Toxins  
; FILE REFERENCE: P-AR 4803  
; CURRENT APPLICATION NUMBER: US/09/942,024  
; CURRENT FILING DATE: 2001-08-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-942-024-21

Query Match 58.6%; Score 95; DB 10; Length 288;  
Best Local Similarity 65.5%; Pred. No. 4.9e-05;  
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSGIOQLERSIRELHDF 30  
Db 188 SKQALSEIETRHSEIITKLENSIRELHDMF 216

RESULT 13  
US-09-942-098-21  
; Sequence 21, Application US/09942098  
; Publication No. US20030143651A1  
; GENERAL INFORMATION:  
; APPLICANT: Steward, Lance E.  
; APPLICANT: Fernandez-Salas, Ester  
; APPLICANT: Aoki, Kei Roger  
; TITLE OF INVENTION: Fret Protease Assays For Clostridial  
; TITLE OF INVENTION: Toxins  
; FILE REFERENCE: P-AR 4802  
; CURRENT APPLICATION NUMBER: US/09/942,098  
; CURRENT FILING DATE: 2001-08-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-942-098-21

Query Match 58.6%; Score 95; DB 10; Length 288;  
Best Local Similarity 65.5%; Pred. No. 4.9e-05;  
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSGIOQLERSIRELHDF 30  
Db 188 SKQALSEIETRHSEIITKLENSIRELHDMF 216

RESULT 14  
US-10-261-161-16  
; Sequence 16, Application US/10261161  
; Publication No. US20040072270A1  
; GENERAL INFORMATION:

; APPLICANT: Fernandez-Salas, Ester  
; APPLICANT: Steward, Lance E.  
; APPLICANT: Aoki, Kei Roger  
; TITLE OF INVENTION: Cell-Based Fluorescence Resonance Energy  
; TITLE OF INVENTION: Transfer (FRET) Assays For Clostridial Toxins  
; FILE REFERENCE: P-AR 4804  
; CURRENT APPLICATION NUMBER: US/10/261,161  
; CURRENT FILING DATE: 2002-09-27  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-261-161-16

Query Match 58.6%; Score 95; DB 15; Length 288;  
Best Local Similarity 65.5%; Pred. No. 4.9e-05;  
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSGIOQLERSIRELHDF 30  
Db 188 SKQALSEIETRHSEIITKLENSIRELHDMF 216

RESULT 15  
US-09-942-024-23  
; Sequence 23, Application US/09942024  
; Publication No. US20030143650A1  
; GENERAL INFORMATION:  
; APPLICANT: Steward, Lance E.  
; APPLICANT: Fernandez-Salas, Ester  
; APPLICANT: Aoki, Kei Roger  
; TITLE OF INVENTION: Fret Protease Assays For Botulinum  
; TITLE OF INVENTION: Serotype A/E Toxins  
; FILE REFERENCE: P-AR 4803  
; CURRENT APPLICATION NUMBER: US/09/942,024  
; CURRENT FILING DATE: 2001-08-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-942-024-23

Query Match 58.0%; Score 94; DB 10; Length 288;  
Best Local Similarity 65.5%; Pred. No. 6.8e-05;  
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSGIOQLERSIRELHDF 30  
Db 188 SKQALSEIETRHSEIITKLENSIRELHDMF 216

Search completed: November 10, 2004, 16:35:45  
Job time : 32.8287 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:36:51 ; Search time 6.6713 seconds  
(without alignments)  
475.942 Million cell updates/sec

Title: US-10-092-750-9

Sequence: 162  
1 ATGALNEISARHSGIQQLERSIRLHDIPTFL 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	150	92.6	297	2	SS2726
2	150	92.6	297	2	I38517
3	143	88.3	298	2	E48213
4	103	63.6	288	2	JN0466
5	103	63.6	289	2	S51193
6	103	63.6	290	2	C48213
7	100	61.7	288	2	JU0136
8	100	61.7	288	2	B48213
9	95	58.6	259	2	G01485
10	95	58.6	288	2	U01615
11	95	58.6	298	2	A38141
12	89	54.9	269	2	I83198
13	89	54.9	283	2	I83197
14	89	54.9	289	2	G01969
15	89	54.9	289	2	I60170
16	89	54.9	289	2	D48213
17	88	54.3	291	2	T37265
18	88	54.3	291	2	T37266
19	78	48.1	287	2	JE0094
20	76.5	47.2	291	2	A58573
21	74	45.7	341	2	T08459
22	72	44.4	305	2	C85042
23	69	42.6	284	2	T41524
24	69	42.6	341	2	T48847
25	66	40.7	298	2	D86246
26	65	40.1	303	2	G96638
27	63	38.9	315	2	T13654
28	62	38.3	255	2	H85198
29	62	38.3	275	2	D71447

30	61	37.7	290	2	S39569	syntaxin-related p
31	58	35.8	295	2	S55130	syntaxin-related p
32	56	34.6	296	2	E83161	probable transcrip
33	55	34.0	280	2	D86218	protein P22013.4 l
34	55	34.0	310	2	T00709	syntaxin-related p
35	55	34.0	331	2	D64650	hypothetical prote
36	55	34.0	331	2	E71854	hypothetical prote
37	54	33.3	96	2	D81091	glutaryl-tRNA (Gln
38	54	33.3	129	2	T17322	hypothetical prote
39	54	33.3	307	2	T50493	syntaxin-like prote
40	54	33.3	875	2	T10340	hypothetical prote
41	53	32.7	141	2	A85994	probable transcrip
42	53	32.7	141	2	E91148	zinc (II) responsi
43	53	32.7	141	2	T67892	zinc (II) responsi
44	53	32.7	275	2	T21862	hypothetical prote
45	53	32.7	448	2	C86447	FSD14.3 protein -

## ALIGNMENTS

## RESULT 1

S52726  
syntaxin-4 - human

C:Species: Homo sapiens (man)

C:Date: 19-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004

C:Accession: S52726

R:Jagadeish, M.N.; Fernandez, C.S.  
submitted to the EMBL Data Library, December 1994

A:Description: Isolation and sequence analysis of the syntaxin-4 encoding region from 1

A:Accession: S52726

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-297 <JAG>

A:Cross-references: UNIPROT:Q12846; EMBL:X85784; NID:G758104; PIDN:CAA59769.1; PID:G758

C:Superfamily: syntaxin

Query Match 92.6%; Score 150; DB 2; Length 297;  
Best Local Similarity 96.9%; Pred. No. 4.4e-12;  
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TROALNEISARHSGIQQLERSIRLHDIPTFL 33

DB 196 TROALNEISARHSGIQQLERSIRLHDIPTFL 227

## RESULT 2

I38517  
syntaxin - human

C:Species: Homo sapiens (man)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004

C:Accession: I38517

R:Li, H.; Hodge, D.R.; Pei, G.X.; Seth, A.

Gene 143, 303-304, 1994

A:Title: Isolation and sequence analysis of the human syntaxin-encoding gene.

A:Reference number: I38517; MUID:94266173; PMID:8206394

A:Accession: I38517

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-297 <RES>

A:Cross-references: UNIPROT:Q12846; EMBL:U07158; NID:G463906; PIDN:AAA20967.1; PID:G46

Query Match 92.6%; Score 150; DB 2; Length 297;  
Best Local Similarity 96.9%; Pred. No. 4.4e-12;  
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TROALNEISARHSGIQQLERSIRLHDIPTFL 33

DB 196 TROALNEISARHSGIQQLERSIRLHDIPTFL 227

RESULT 3  
E48213  
synixin 4 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #ext\_change 09-Jul-2004  
C:Accession: E48213  
R:Bennett, M.K.; Garcia-Arreaga, J.E.; Elferink, L.A.; Peterson, K.; Fleming, A.M.; Hazu  
Cell 74, 863-873, 1993  
A:Title: The synixin family of vesicular transport receptors.  
A:Reference number: A48213; MUID:93386759; PMID:7690687  
A:Accession: E48213  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-298 <BEN>  
A:Cross-references: UNIPROT:Q08850; GB:L20821; NID:G349320; PIDN:AAA03046.1; PID:G349321  
C:Superfamily: synixin

Query Match 88.3%; Score 143; DB 2; Length 298;  
Best Local Similarity 90.6%; Pred. No. 3.6e-11;  
Matches 29; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSIQQLERSIRELHDIIF 33  
Db 196 TROALNEISARHSIQQLERSIRELHDIIF 227

RESULT 4  
JN0466  
epimorphin - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #ext\_change 04-Sep-1998  
C:Accession: JN0466  
R:Hirai, Y.  
Biochem. Res. Commun. 191, 1332-1337, 1993  
A:Title: Molecular cloning of human epimorphin: identification of isoforms and their uni  
A:Reference number: JN0466; MUID:93221506; PMID:8465509  
A:Accession: JN0466  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-288 <HIR>  
A:Experimental source: placenta  
C:Genetics:  
A:Gene: GDB:EPIM  
A:Cross-references: GDB:362733; OMIM:132350  
A:Map position: 7q-7q  
C:Superfamily: synixin  
F:265-288/Domain: transmembrane #status predicted <TM>

Query Match 63.6%; Score 103; DB 2; Length 288;  
Best Local Similarity 69.0%; Pred. No. 5.8e-06;  
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSIQQLERSIRELHDIIF 30  
Db 187 TROALNEISARHSIQQLERSIRELHDIIF 215

RESULT 5  
S51193  
epimorphin - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 01-Aug-1995 #sequence\_revision 07-Jun-1996 #ext\_change 09-Jul-2004  
C:Accession: A38216; S51193  
R:Hirai, Y.; Takebe, K.; Takashina, M.; Kobayashi, S.; Takeichi, M.  
Cell 69, 471-481, 1992  
A:Title: Epimorphin: a mesenchymal protein essential for epithelial morphogenesis.  
A:Reference number: A38216; MUID:92257588; PMID:1581962  
A:Accession: A38216  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-289 <HIR>  
A:Cross-references: UNIPROT:Q00262; GB:D10475; NID:G220402; PIDN:BA01278.1; PID:G220403  
A:Note: sequence extracted from NCBI backbone (NCBIN:101770, NCBI:101772)

R:Hirai, Y.  
Eur. J. Biochem. 225, 1133-1139, 1994  
A:Title: Sodium-dodecyl-sulfate-resistant complex formation of epimorphin monomers and  
A:Reference number: S51193; MUID:95045543; PMID:7957204  
A:Accession: S51193  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-12, 'I', '14-28, 'P', '30-78, 'N', '80-135, 'F', '137-154, 'T', '156-289 <HI2>  
C:Superfamily: synixin  
C:Keywords: membrane trafficking

Query Match 63.6%; Score 103; DB 2; Length 289;  
Best Local Similarity 69.0%; Pred. No. 5.8e-06;  
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSIQQLERSIRELHDIIF 30  
Db 188 TROALNEISARHSIQQLERSIRELHDIIF 216

RESULT 6  
C48213  
synixin 2 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #ext\_change 07-May-1999  
C:Accession: C48213  
R:Bennett, M.K.; Garcia-Arreaga, J.E.; Elferink, L.A.; Peterson, K.; Fleming, A.M.; Hazu  
Cell 74, 863-873, 1993  
A:Title: The synixin family of vesicular transport receptors.  
A:Reference number: A48213; MUID:93386759; PMID:7690687  
A:Accession: C48213  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-290 <BEN>  
A:Cross-references: GB:L20823  
A:Note: authors translation is shown for the codon ACG at position 287  
C:Superfamily: synixin

Query Match 63.6%; Score 103; DB 2; Length 290;  
Best Local Similarity 69.0%; Pred. No. 5.8e-06;  
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TROALNEISARHSIQQLERSIRELHDIIF 30  
Db 189 TROALNEISARHSIQQLERSIRELHDIIF 217

RESULT 7  
JU0136  
synaptocanalin I - bovine  
N:Alternate names: SNAP receptor  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #ext\_change 09-Jul-2004  
C:Accession: JU0136; S32361  
R:Morita, T.; Mori, H.; Sakimura, K.; Mishina, M.; Sekine, Y.; Tsugita, A.; Odani, S.;  
Biomed. Res. 13, 357-364, 1992  
A:Title: Synaptocanalin I, a protein associated with brain omega-conotoxin-sensitive ca  
A:Reference number: JU0136  
A:Accession: JU0136  
A:Molecule type: mRNA  
A:Residues: 1-288 <MOR>  
A:Cross-references: UNIPROT:P61267  
R:Soellner, T.; Whiteheart, S.W.; Brunner, M.; Erdjument-Bromage, H.; Geromano, S.; T  
Nature 362, 318-324, 1993  
A:Title: SNAP receptors implicated in vesicle targeting and fusion.  
A:Reference number: S32360; MUID:93205116; PMID:8455717  
A:Accession: S32361  
A:Status: preliminary  
A:Molecule type: Protein  
A:Residues: 46-55, 'X', '57, 'X', '59-78, 'A', '80-81, '125-137, '158-181 <SOE>  
C:Superfamily: synixin  
C:Keywords: transmembrane protein  
F:165-288/Domain: transmembrane #status predicted <TM>



```
Query Match      61.7%; Score 100; DB 2; Length 288;
Best Local Similarity 69.0%; Pred. No. 1.4e-05;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSIGQQLERSIRELHDF 30
       187 TKQALNEIETRNHEIKLTSTIRELHDMF 215

RESULT 8
B48213
Syntaxin 1B - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: B48213; S33605
R:Bennett, M.K.; Garcia-Arraras, J.E.; Elferink, L.A.; Peterson, K.; Fleming, A.M.; Hazu
Cell 74, 863-873, 1993
A:Title: The syntaxin family of vesicular transport receptors.
A:Reference number: A48213; MUID:93386759; PMID:7690687
A:Accession: B48213
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-288 <BEN>
A:Cross-references: UNIPROT:P61265; GB:M95735; NID:g207138; PIDN:AAA42197.1; PID:g207139
Science 257, 255-259, 1992
A:Title: Syntaxin: a synaptic protein implicated in docking of synaptic vesicles at pres
A:Reference number: S33604; MUID:92335866; PMID:1321498
A:Accession: S33605
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-288 <BEN2>
A:Cross-references: EMBL:M95735; NID:g207138; PIDN:AAA42197.1; PID:g207139
C:Superfamily: syntaxin
C:Keywords: transmembrane protein

Query Match      61.7%; Score 100; DB 2; Length 288;
Best Local Similarity 69.0%; Pred. No. 1.4e-05;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSIGQQLERSIRELHDF 30
       187 TKQALNEIETRNHEIKLTSTIRELHDMF 215

RESULT 9
G01485
Syntaxin - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C:Accession: G01485
R:Zhang, R.
Submitted to the EMBL Data Library, August 1994
A:Reference number: G07331
A:Accession: G01485
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-259 <ZHA>
A:Cross-references: EMBL:U12918; NID:g531457; PIDN:AAA20940.1; PID:g531458
C:Superfamily: syntaxin

Query Match      58.6%; Score 95; DB 2; Length 259;
Best Local Similarity 65.5%; Pred. No. 5.7e-05;
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSIGQQLERSIRELHDF 30
       159 SKQALSEIETRHSRIIKLENSIRELHDMF 167

RESULT 10
Q01615
neuron-specific antigen HPC-1 - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C:Accession: Q01615
R:Inoue, A.; Akagawa, K.
Biochem. Biophys. Res. Commun. 187, 1144-1150, 1992
A:Title: Neuron-specific antigen HPC-1 from bovine brain reveals strong homology to ep
A:Reference number: Q01615; MUID:92412081; PMID:1530610
A:Accession: Q01615
A:Molecule type: mRNA
A:Residues: 1-288 <INO>
A:Cross-references: UNIPROT:P32850
A:Experimental source: brain
C:Keywords: membrane protein; membrane trafficking

Query Match      58.6%; Score 95; DB 2; Length 288;
Best Local Similarity 65.5%; Pred. No. 6.4e-05;
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSIGQQLERSIRELHDF 30
       188 SKQALSEIETRHSRIIKLENSIRELHDMF 216

RESULT 11
A38141
neuronal cell membrane antigen HPC-1 - rat (fragment)
N:Alternate names: synaptotagmin associated 35kDa protein; syntaxin 1A
C:Species: Rattus norvegicus (Norway rat)
C>Date: 04-Mar-1993 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C:Accession: A38141; A48213; S33604; I55372; J53360; PC413
R:Inoue, A.; Obara, K.; Akagawa, K.
T. Biol. Chem. 267, 10613-10619, 1992
A:Title: Cloning and sequence analysis of cDNA for a neuronal cell membrane antigen, H
A:Reference number: A38141; MUID:92268107; PMID:1587842
A:Accession: A38141
A:Molecule type: mRNA
A:Residues: 1-298 <INO>
A:Cross-references: UNIPROT:P32851; GB:D10392; NID:g220776; PIDN:BAA01231.1; PID:g22077
A:Experimental source: hippocampus
A>Note: sequence extracted from NCBI backbone (NCBI:103808, NCBI:103809)
R:Bennett, M.K.; Garcia-Arraras, J.E.; Elferink, L.A.; Peterson, K.; Fleming, A.M.; Haz
Cell 74, 863-873, 1993
A:Title: The syntaxin family of vesicular transport receptors.
A:Reference number: A48213; MUID:93386759; PMID:7690687
A:Accession: A48213
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 11-298 <BEN>
A:Cross-references: GB:M95734
R:Bennett, M.K.; Calakos, N.; Scheller, R.H.
Science 257, 255-259, 1992
A:Title: Syntaxin: a synaptic protein implicated in docking of synaptic vesicles at pre
A:Reference number: S33604; MUID:92335866; PMID:1321498
A:Accession: S33604
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 14-298 <BE2>
A:Cross-references: EMBL:M95734; NID:g207126; PIDN:AAA42195.1; PID:g207127
R:Yoshida, A.; Oho, C.; Omori, A.; Kuwahara, R.; Ito, T.; Takahashi, M.
U. Biol. Chem. 267, 24925-24928, 1992
A:Title: HPC-1 is associated with synaptotagmin and omega-conotoxin receptor.
A:Reference number: I55372; MUID:93094187; PMID:1334074
A:Accession: I55372
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 14-298 <RES>
A:Cross-references: GB:D12519; NID:g220901; PIDN:BAA02089.1; PID:g220902
A:Experimental source: brain
A>Note: parts of this sequence, including the amino end of the mature protein, were det
R:Fujiwara, T.; Yamamori, T.; Yamaguchi, K.; Akagawa, K.
Biochem. Biophys. Res. Commun. 231, 352-355, 1997
```

```

A>Title: Interaction of HPC-1/syntaxin 1A with the cytoskeletal protein, tubulin.
A/Reference number: JCS360; MUID:97223456; PMID:9070277
A/Accession: JCS360
A/Molecule type: protein
A/Residues: 11-130 <FUD>
A/Accession: PC4313
A/Molecule type: protein
A/Residues: 54-68/79-97/99-116 <FU2>
C/Comment: This protein is involved in morphological change, possibly via interaction
C/Genes: SAP
A/Superfamily: syntaxin
C/Keywords: membrane trafficking; surface antigen; transmembrane protein
F/99-116/Domain: tublin binding #status predicted <TB>
F/176-298/Domain: transmembrane #status predicted <TM>

Query Match      58.6%; Score 95; DB 2; Length 298;
Best Local Similarity 65.5%; Pred. No. 6.6e-05;
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDF 30
Db      198 SKQALSEIRGRKDIVRLSSIKELHDMF 226

RESULT 12
183198
Syntaxin 3C - mouse
C/Species: Mus musculus (house mouse)
C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C/Accession: 183198
R/Ibaraki, K.; Horikawa, H.P.; Morita, T.; Mori, H.; Sakimura, K.; Mishina, M.; Saisu, H.
Biochem. Biophys. Res. Commun. 211, 997-1005, 1995
A/Title: Identification of four different forms of syntaxin 3.
A/Reference number: 160170; MUID:95321968; PMID:7598732
A/Accession: 183198
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-269 <RES>
A/Cross-references: UNIPROT:Q64704; GB:D29799; NID:9924271; PIDN:BA06182.1; PID:9924272
C/Superfamily: syntaxin

Query Match      54.9%; Score 89; DB 2; Length 269;
Best Local Similarity 58.6%; Pred. No. 0.00036;
Matches 17; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDF 30
Db      169 SKQALSEIRGRKDIVRLSSIKELHDMF 197

RESULT 13
183197
Syntaxin 3B - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C/Accession: 183197
R/Ibaraki, K.; Horikawa, H.P.; Morita, T.; Mori, H.; Sakimura, K.; Mishina, M.; Saisu, H.
Biochem. Biophys. Res. Commun. 211, 997-1005, 1995
A/Title: Identification of four different forms of syntaxin 3.
A/Reference number: 160170; MUID:95321968; PMID:7598732
A/Accession: 183197
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-283 <RES>
A/Cross-references: UNIPROT:Q64704; GB:D29798; NID:9924269; PIDN:BA06181.1; PID:9924270
C/Superfamily: syntaxin

Query Match      54.9%; Score 89; DB 2; Length 283;
Best Local Similarity 58.6%; Pred. No. 0.00038;
Matches 17; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDF 30

```

```

Db      187 SKQALSEIRGRKDIVRLSSIKELHDMF 215

RESULT 14
G01969
Syntaxin 3 - human
C/Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C/Accession: G01969
R/Naren, A.P.; Bradbury, N.A.; Bennett, M.K.; Kirk, K.L.
submitted to the EMBL Data Library, July 1995
A/Reference number: G08957
A/Accession: G01969
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-289 <NAR>
A/Cross-references: UNIPROT:Q13277; EMBL:U32315; NID:9929930; PIDN:AAA75303.1; PID:99299
C/Superfamily: syntaxin

Query Match      54.9%; Score 89; DB 2; Length 289;
Best Local Similarity 58.6%; Pred. No. 0.00039;
Matches 17; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDF 30
Db      187 SKQALSEIRGRKDIVRLSSIKELHDMF 215

RESULT 15
160170
Syntaxin 3A - mouse
C/Species: Mus musculus (house mouse)
C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C/Accession: 160170
R/Ibaraki, K.; Horikawa, H.P.; Morita, T.; Mori, H.; Sakimura, K.; Mishina, M.; Saisu, H.
Biochem. Biophys. Res. Commun. 211, 997-1005, 1995
A/Title: Identification of four different forms of syntaxin 3.
A/Reference number: 160170; MUID:95321968; PMID:7598732
A/Accession: 160170
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-289 <RES>
A/Cross-references: UNIPROT:Q64704; GB:D29797; NID:9924267; PIDN:BA06180.1; PID:9924268
C/Superfamily: syntaxin

Query Match      54.9%; Score 89; DB 2; Length 289;
Best Local Similarity 58.6%; Pred. No. 0.00039;
Matches 17; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      2 TROALNEISARHSGIOQLERSIRELHDF 30
Db      187 SKQALSEIRGRKDIVRLSSIKELHDMF 215

Search completed: November 10, 2004, 12:29:10
Job time : 7.6713 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:27:01 ; Search time 36.1065 Seconds  
(without alignments)  
525.871 Million cell updates/sec

Title: US-10-092-750-9

Perfect score: 162

Sequence: 1 ATRQALNEISARHSIGIQLEERSIRELHDIPTFL 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : UniProt\_02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	92.6	297	1	STX4_HUMAN
2	150	92.6	297	2	AAP35990
3	146	90.1	298	1	STX4_MOUSE
4	146	90.1	298	2	Q80WT8
5	143	88.3	298	1	STX4_RAT
6	116	71.6	297	2	Q7ZV18
7	116	71.6	297	2	AAC94577
8	104	64.2	96	2	Q9W788
9	103	63.6	237	2	Q7E577
10	103	63.6	286	2	Q8J1P4
11	103	63.6	288	1	EPMO_HUMAN
12	103	63.6	289	1	EPMO_MOUSE
13	103	63.6	289	2	Q80W45
14	103	63.6	290	1	EPMO_RAT
15	101	62.3	141	2	Q6DFT1
16	101	62.3	288	2	Q919P6
17	101	62.3	290	2	Q6DD19
18	100	61.7	137	2	Q7T1K3
19	100	61.7	288	1	ST1B_HUMAN
20	100	61.7	288	1	ST1C_BOVIN
21	100	61.7	288	1	ST1C_HUMAN
22	100	61.7	288	1	ST1C_MOUSE
23	100	61.7	288	1	ST1C_RAT
24	100	61.7	288	1	ST1C_SHEEP
25	98	60.5	282	2	O42340
26	96	59.3	288	2	Q7SYFO
27	96	59.3	288	2	AAP83589
28	95	58.6	251	2	AAS07470
29	95	58.6	250	2	Q7ZSK3
30	95	58.6	288	1	ST1A_BOVIN
31	95	58.6	288	1	ST1A_HUMAN

32	95	58.6	288	1	ST1A_RAT	P32851 rattus norv
33	95	58.6	288	2	Q9N0X9	Q9N0X9 macaca mla
34	95	58.6	288	2	AAS07469	AAS07469 homo sapi
35	95	58.6	288	2	AAB64644	AAB64644 homo sapi
36	95	58.6	288	2	CAG33299	CAG33299 homo sapi
37	94	58.0	287	2	AAB47133	AAB47133 mus muscu
38	94	58.0	288	1	ST1A_MOUSE	Q35526 mus muscu
39	94	58.0	288	2	Q35535	Q35535 mus muscu
40	93	57.4	221	2	Q6DEJ2	Q6DEJ2 brachydanio
41	93	57.4	288	2	Q6PHV0	Q6PHV0 brachydanio
42	93	57.4	288	2	AAB56333	AAB56333 brachydan
43	90	55.6	286	2	Q6GLJ9	Q6GLJ9 xenopus lae
44	89	54.9	258	2	Q7Z338	Q7Z338 homo sapien
45	89	54.9	287	2	Q8CAN1	Q8CAN1 mus muscu

#### ALIGNMENTS

RESULT 1  
ID STX4\_HUMAN STANDARD: PRT: 297 AA.  
AC Q12846; Q15525;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Syntaxin 4.  
GN Name=STX4A; Synonyms=STX4;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=94266173; PubMed=8206394;  
RA Li H., Hodge D.R., Pei G.K., Seth A.;  
RT "Isolation and sequence analysis of the human syntaxin-encoding  
RT gene.";  
RL Gene 143:303-304(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=96332494; PubMed=8760387;  
RA Jagadeish M.N., Fernandez C.S., Hewish D.R., Macaulay S.L., Gough K.H.,  
RA Grusovin J., Verkuylen A., Cosgrove L., Alafaci A., Frenkel M.J.,  
RA Ward C.W.;  
RT "Insulin-responsive tissues contain the core complex protein SNAP-25  
RT (synaptosomal-associated protein 25) A and B isoforms in addition to  
RT syntaxin 4 and synaptobrevins 1 and 2.";  
RL Biochem. J. 317:945-954(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Peripheral blood neutrophils;  
RX Napolkina S., Lazo P.A., Mollinedo F.;  
RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lens epithelium;  
RX Rae J.U., Shepard A.R.;  
RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=23368257; PubMed=12477933; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner J., Sherman C.W., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Sabetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Hale S., Garcia A.M., Gunatane P.H.,  
 RA Richards S., Morley K.C., Hsieh S., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Phelan M., Heiton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Batteigeld J.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schenker A., Schein J.E., Jones S.J.M., Maira M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Potentially involved in docking of synaptic vesicles at  
 CC presynaptic active zones.  
 CC -1- SUBUNIT: Binds STXB3 and STXB6. Interacts with SNAP23 and  
 CC SNAP25BP (By similarity). Binding to STXB3 excludes binding with  
 CC SNAP25.  
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (Potential).  
 CC -1- SIMILARITY: Belongs to the syntaxin/epimorphin family.  
 CC -1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U07158; AAA20967.1; -;  
 DR EMBL: X85784; CAA59769.1; -;  
 DR EMBL: AJ000541; CAA04174.1; -;  
 DR EMBL: AF026007; AAB8810.1; -;  
 DR EMBL: AF18489; AAG40313.1; -;  
 DR EMBL: BC002435; AA02436.1; -;  
 DR PIR: I38517; I38517.  
 DR PIR: S52726; S52726.  
 DR HSP: P32851; I07H.  
 DR OGP: Q12846; -;  
 DR Genew: HGNC:11439; STX4A.  
 DR MIM: 186591; -;  
 DR GO: GO:0016020; C-membrane; TAS.  
 DR InterPro: IPR006012; Syntaxin.  
 DR InterPro: IPR006011; Syntaxin\_N.  
 DR InterPro: IPR010989; t-snare.  
 DR InterPro: IPR00727; t-SNARE.  
 DR Pfam: PF05739; SNARE; 1.  
 DR SMART: SM00804; Syntaxin; 1.  
 DR SMART: SM00503; SYN: 1.  
 DR SMART: SM00397; t-SNARE; 1.  
 DR PROSITE: PS00914; SYNTAXIN; 1.  
 DR PROSITE: PS50192; t-SNARE; 1.  
 KW Coiled coil; Neurotransmitter transport; Transmembrane.  
 FT DOMAIN 1 275 Cytoplasmic (Potential).  
 FT TRANSMEM 276 296 Anchored for type IV membrane protein  
 FT (Potential)  
 FT DOMAIN 297 297 Extracellular (Potential).  
 FT DOMAIN 297 297 Coiled coil (Potential).  
 FT DOMAIN 297 297 Coiled coil (Potential).  
 FT DOMAIN 297 297 t-SNARE coiled-coil homology.  
 FT CONFLICT 174 174 E -> D (in Ref. 1).  
 FT CONFLICT 269 269 A -> V (in Ref. 1).  
 SQ SEQUENCE 297 AA; 34180 MW; 5084FD1C49A86BAA CRC64;

Db 196 TROALNEISARHSGICQLERSIRELHDIFTL 227  
 RESULT 2  
 AAP35990 PRELIMINARY; PRT; 297 AA.  
 AC AAP35990;  
 DT 02-MAR-2004 (TRMBLrel. 27, Created)  
 DT 02-MAR-2004 (TRMBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TRMBLrel. 27, Last annotation update)  
 DE Syntaxin 4A (Placental).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 ON NCBI\_Taxid=9606;  
 RX SEQUENCE FROM N.A.  
 RA Kaine N., Chen X., Rolfe A., Halleck A., Hines L., Eisenstein S.,  
 RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,  
 RA Phelan M., Farmer A.,  
 RT "Cloning of human full-length cDNAs in BD Creator (TM) System Donor  
 RT vector.";  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BT007326; AAP35990.1; -;  
 SQ SEQUENCE 297 AA; 34180 MW; 5084FD1C49A86BAA CRC64;

Query Match 92.6%; Score 150; DB 2; Length 297;  
 Best Local Similarity 96.9%; Pred. No. 2e-11;  
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 TROALNEISARHSGICQLERSIRELHDIFTL 33  
 Db 196 TROALNEISARHSGICQLERSIRELHDIFTL 227

RESULT 3  
 STX4 MOUSE STANDARD; PRT; 298 AA.  
 ID STX4 MOUSE  
 AC P70452;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Syntaxin 4.  
 GN Name=Stx4a; Synonyms=Stx4;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_Taxid=10090;  
 RX SEQUENCE FROM N.A.  
 RP MEDLINE=97197781; PubMed=9045631;  
 RA Tellam J.T., Macaulay S.L., McIntosh S., Hewish D.R., Ward C.W.,  
 RA James D.E.;  
 RT "Characterization of Munc-18c and syntaxin-4 in 3T3-L1 adipocytes.  
 RT Putative role in insulin-dependent movement of GLUT-4.";  
 RL J. Biol. Chem. 272:6179-6186(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schuetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Hale S., Garcia A.M., Gunatane P.H.,  
 RA Richards S., Morley K.C., Hsieh S., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Phelan M., Heiton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Fahey J., Heiton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schencher A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP INTERACTION WITH SNAP23 AND SNAP25BP.  
 RX PubMed=12877659; DOI=10.1042/BJ20030427;  
 RA Buxton P., Zhang X.-M., Walsh B., Sritanana A., Schenberger I.,  
 RA Manickam E., Rowe T.;  
 RT "Identification and characterization of Synapin as a ubiquitously  
 RT expressed SNARE-binding protein that interacts with SNAP23 in non-  
 RT neuronal cells";  
 RL Biochem. J. 375:433-440(2003).  
 CC -1- FUNCTION: Potentially involved in docking of synaptic vesicles at  
 CC presynaptic active zones.  
 CC SUBUNIT: Binds syntaxin and syntaxin6 (By similarity). Interacts with  
 CC SNAP23 and SNAP25BP. Binding to syntaxin6 excludes binding with  
 CC SNAP25 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (Potential).  
 CC -1- SIMILARITY: Belongs to the syntaxin/epimorphin family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: U76832; AAB18991.1; -;  
 DR EMBL: BC005791; AAH05791.1; -;  
 DR EMBL: BC011491; AAH11491.1; -;  
 DR HSSP: P32851; 1JTH.  
 DR MGD: MGI:893577; Stx4a.  
 DR GO: GO:0005515; F:protein binding; IPI.  
 DR InterPro: IPR006012; Syntaxin.  
 DR InterPro: IPR006011; Syntaxin\_N.  
 DR InterPro: IPR010989; t-snare.  
 DR InterPro: IPR000727; t-SNARE.  
 DR Pfam: PF05739; SNARE; 1.  
 DR Pfam: PF0804; Syntaxin; 1.  
 DR SMART: SM00503; SYN; 1.  
 DR SMART: SM00397; t-SNARE; 1.  
 DR PROSITE: PS00914; SYNTAXIN; 1.  
 DR PROSITE: PS50192; t-SNARE; 1.  
 KW Coiled coil; Neurotransmitter transport; Transmembrane.  
 FT DOMAIN 1 274 Cytoplasmic (Potential).  
 FT TRANSMEM 275 295 Anchor for type IV membrane protein  
 FT DOMAIN 296 298 Extracellular (Potential).  
 FT DOMAIN 38 163 Coiled coil (Potential).  
 FT DOMAIN 200 262 t-SNARE coiled-coil homology.  
 SQ SEQUENCE 298 AA; 34165 MW; FCD147E1126CEC1 CRC64;  
 Query Match 90.1%; Score 146; DB 1; Length 298;  
 Best Local Similarity 93.8%; Pred. No. 6,6e-11;  
 Matches 30; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)  
 DE Syntaxin 4A (Placental).  
 GN Name=Stx4a;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Heien F.,  
 RA Diatchenko L., Marzella K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueclin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,  
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Heltan E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalhus D.E., Schencher A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the syntaxin/epimorphin family.  
 DR EMBL: BC052023; AAH52023.1; -;  
 DR HSSP: P32851; 1JTH.  
 DR MGD: MGI:893577; Stx4a.  
 DR GO: GO:0005515; F:protein binding; IPI.  
 DR InterPro: IPR006012; Syntaxin.  
 DR InterPro: IPR006011; Syntaxin\_N.  
 DR InterPro: IPR010989; t-snare.  
 DR InterPro: IPR000727; t-SNARE.  
 DR Pfam: PF05739; SNARE; 1.  
 DR Pfam: PF0804; Syntaxin; 1.  
 DR SMART: SM00503; SYN; 1.  
 DR SMART: SM00397; t-SNARE; 1.  
 DR PROSITE: PS00914; SYNTAXIN; 1.  
 DR PROSITE: PS50192; t-SNARE; 1.  
 SQ SEQUENCE 298 AA; 34177 MW; 65D147793826CED CRC64;  
 Query Match 90.1%; Score 146; DB 2; Length 298;  
 Best Local Similarity 93.8%; Pred. No. 6,6e-11;  
 Matches 30; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 5  
 ID STX4\_RAT STANDARD; PRT; 298 AA.  
 AC Q08850;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE Syntaxin 4.  
 GN Name=Stx4a; Synonyms=Stx4;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93386759; PubMed=7690687;  
 RA Bennett M.K., Garcia-Arriaza J.B., Elferink L.A., Peterson K.E.,  
 RA Fleming A.M., Hazuka C.D., Scheller R.H.,  
 RT "The syntaxin family of vesicular transport receptors."  
 RL Cell 74:863-873 (1993).  
 RN [2]  
 RP INTERACTION WITH STXBP6.  
 RX MEDLINE=22140381; PubMed=12145319; DOI=10.1074/jbc.M204929200;  
 RA Scates S.J., Hesser B.A., Masuda E.S., Scheller R.H.,  
 RT "Syntaxin, a novel syntaxin-binding protein that may regulate SNARE  
 RT complex assembly."  
 RL J. Biol. Chem. 277:28271-28279 (2002).  
 CC -1- FUNCTION: Potentially involved in docking of synaptic vesicles at  
 CC presynaptic active zones.  
 CC -1- SUBUNIT: Interacts with SNAP23 and SNAP25BP (By similarity). Bands  
 CC STXBP3 and STXBP6 (By similarity). Binding to STXBP3 excludes  
 CC binding with SNAP25 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (Potential).  
 CC -1- TISSUE SPECIFICITY: Heart, spleen, skeletal muscle and kidney.  
 CC -1- SIMILARITY: Belongs to the syntaxin/epimorphin family.  
 CC -1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: L20821; AAA03046.1; -.  
 DR F1R; E48213; E48213.  
 DR HSP; P32851; IUTR.  
 DR RD; 621019; Stx4a.  
 DR InterPro: IPR006012; Syntaxin.  
 DR InterPro: IPR006011; Syntaxin\_N.  
 DR InterPro: IPR010989; t-snare.  
 DR InterPro: IPR007027; t-SNARE.  
 DR Pfam: PF05739; SNARE; 1.  
 DR Pfam: PF00804; Syntaxin; 1.  
 DR PROSITE: PS00914; SYNTAXIN; 1.  
 DR PROSITE: PS0192; t-SNARE; 1.  
 DR Coiled coil; Neurotransmitter transport; Transmembrane.  
 FT DOMAIN 1 274 Cytoplasmic (Potential).  
 FT TRANSMEM 275 295 Anchor for type IV membrane protein  
 FT DOMAIN 296 298 (Potential).  
 FT DOMAIN 38 163 Extracellular (Potential).  
 FT DOMAIN 200 262 Coiled coil (Potential).  
 FT DOMAIN 298 298 t-SNARE coiled-coil homology.  
 SQ SEQUENCE 298 AA; 34209 MW; 9E854270DFB3C96 CRC64;  
 Query Match 88.3%; Score 143; DB 1; Length 298;  
 Best Local Similarity 90.6%; Pred. No. 1.6e-10;  
 Matches 29; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 TROALNHSARHSGIOLESIRELHDIFTL 33  
 Db 196 TROALNHSARHSGIOLESIRELHDIFTL 227  
 RESULT 6  
 Q7ZV18 PRELIMINARY; PRT; 297 AA.  
 ID Q7ZV18  
 AC Q7ZV18  
 DT 01-JUN-2003 (TReMBurel. 24, Created)  
 DT 01-JUN-2003 (TReMBurel. 24, Last sequence update)  
 DT 01-OCT-2004 (TReMBurel. 28, Last annotation update)  
 DE Placental syntaxin 4A (Placental).  
 GN Name=zgc:56272;

OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole body;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins L., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,  
 RA Ditzchenko L., Marziska K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton W., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein W.J., Ueda T.B., Tomiyaki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulvaney S.J.,  
 RA Bock S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalek U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole body;  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney matrix;  
 RA Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,  
 RA Shang Y., Chen Y., Ruan Y., Jiang C.L., Fan H.Y., Zou L.I.,  
 RA Kanki J.P., Look A.T., Chen Z.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the syntaxin/epimorphin family.  
 CC EMBL: BC046039; AAH46039.1; -.  
 CC EMBL: AY394950; AAQ94577.1; -.  
 DR HSP; P32851; IUTR.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0008565; F:protein transporter activity; IEA.  
 DR GO: GO:006886; P:intracellular protein transport; IEA.  
 DR InterPro: IPR006012; Syntaxin.  
 DR InterPro: IPR006011; Syntaxin\_N.  
 DR InterPro: IPR010989; t-snare.  
 DR InterPro: IPR007027; t-SNARE.  
 DR Pfam: PF05739; SNARE; 1.  
 DR Pfam: PF00804; Syntaxin; 1.  
 DR PROSITE: PS00914; SYNTAXIN; 1.  
 DR PROSITE: PS0192; t-SNARE; 1.  
 SQ SEQUENCE 297 AA; 33920 MW; 96054B62677FCB2 CRC64;  
 Query Match 71.6%; Score 116; DB 2; Length 297;  
 Best Local Similarity 69.7%; Pred. No. 5.9e-07;  
 Matches 23; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 ATROALNHSARHSGIOLESIRELHDIFTL 33  
 Db 196 ATROALNHSARHSGIOLESIRELHDIFTL 228  
 RESULT 7  
 AAQ94577 PRELIMINARY; PRT; 297 AA.  
 ID AAQ94577  
 AC AAQ94577  
 DT 02-MAR-2004 (TReMBurel. 27, Created)  
 DT 02-MAR-2004 (TReMBurel. 27, Last sequence update)

DT 02-MAR-2004 (Tremblrel. 27, last annotation update)  
 DE Placental syntaxin 4A.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Kidney marrow;  
 RA Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,  
 RA Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zou L.I.,  
 RA Kanki J.P., Look A.T., Chen Z.;  
 RT "Gene Expression Profiling in the Zebrafish Kidney Marrow Tissue."  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY394950; AAQ94577.1; -  
 SQ SEQUENCE 297 AA; 33920 MW; 96054B62677FFCB2 CRC64;

Query Match 71.6%; Score 116; DB 2; Length 297;  
 Best Local Similarity 69.7%; Pred. No. 5.9e-07;  
 Matches 23; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGALNIEISARHSGIQQLERSIRELHDIPTFL 33  
 DB 196 ATGALNIEISRHDEIKLERSIKELHDMFOYL 228

RESULT 8  
 Q9W788 PRELIMINARY; PRT; 96 AA.  
 AC Q9W788;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)  
 DE Syntaxin 1 (Fragment).  
 OS Ctenophorus ornatus (Ornate dragon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidossauria; Squamata; Iguania; Acrodonta; Agamidae; Amphibolurinae;  
 OC Ctenophorus.  
 OC NCBI\_TaxID=95347;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Eye;  
 RA Rodger J., Chen P.B., Dunlop S.A., Beazley L.D.;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF146056; AAD37117.1; -  
 DR InterPro: IPR010989; t-snare.  
 DR InterPro: IPR000727; t-snare.  
 DR Pfam: PF05739; SNARE; 1.  
 DR PROSITE: P550192; T-SNARE; 1.  
 FT NON\_TER 1  
 FT NON\_TER 96  
 SQ SEQUENCE 96 AA; 11044 MW; A2F437C508003E24 CRC64;

Query Match 64.2%; Score 104; DB 2; Length 96;  
 Best Local Similarity 72.4%; Pred. No. 7.1e-06;  
 Matches 21; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 TROALNIEISARHSGIQQLERSIRELHDIPTFL 30  
 DB 61 TROALNIEISRHDEIKLERSIKELHDMFOYL 89

RESULT 9  
 Q7TS57 PRELIMINARY; PRT; 237 AA.  
 AC Q7TS57;  
 DT 01-OCT-2003 (Tremblrel. 25, Created)  
 DT 01-OCT-2003 (Tremblrel. 25, last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, last annotation update)  
 DE Syntaxin 2D.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=20033589; PubMed=10564647;  
 RA Quinones B., Riento K., Oikarinen V.M., Hardy S., Bennett M.K.;  
 RA "Syntaxin 2 splice variants exhibit differential expression patterns,  
 RT biochemical properties and subcellular localizations."  
 RL J. Cell Sci. 112:4231-4304 (1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=22623391; PubMed=12737809;  
 RA Low S.H., Li X., Mura M., Kudo N., Quinones B., Weimbs T.;  
 RT "Syntaxin 2 and endobrevin are required for the terminal step of  
 RT cytokinesis in mammalian cells."  
 RL Dev. Cell 4:753-759 (2003).  
 DR EMBL; AY302700; AAF69908.1; -  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR InterPro: IPR006011; Syntaxin\_N.  
 DR InterPro: IPR000727; T-SNARE.  
 DR Pfam: PF05739; SNARE; 1.  
 DR Pfam: PF00804; Syntaxin; 1.  
 DR SMART; SM00503; SYN; 1.  
 DR PROSITE; P550192; T-SNARE; 1.  
 SQ SEQUENCE 237 AA; 27379 MW; 24BD23B463EAAE8 CRC64;

Query Match 63.6%; Score 103; DB 2; Length 237;  
 Best Local Similarity 69.0%; Pred. No. 2.4e-05;  
 Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 TROALNIEISARHSGIQQLERSIRELHDIPTFL 30  
 DB 189 TROALNIEISRHDEIKLERSIKELHDMFOYL 217

RESULT 10  
 Q8UIP4 PRELIMINARY; PRT; 286 AA.  
 AC Q8UIP4;  
 DT 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)  
 DE Coturnix.  
 OS Coturnix coturnix (Common quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix.  
 OC NCBI\_TaxID=9091;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RA Oka Y.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1-SIMILARITY: Belongs to the syntaxin/epimorphin family.  
 DR EMBL; AB076670; BAC00814.1; -  
 DR HSSP; P32851; ISFC.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0008565; P:protein transporter activity; IEA.  
 DR GO: GO:0006886; P:intracellular protein transport; IEA.  
 DR InterPro: IPR006012; Syntaxin.  
 DR InterPro: IPR006011; Syntaxin\_N.  
 DR InterPro: IPR010989; t-snare.  
 DR InterPro: IPR000727; T-SNARE.  
 DR Pfam: PF05739; SNARE; 1.  
 DR Pfam: PF00804; Syntaxin; 1.  
 DR SMART; SM00503; SYN; 1.  
 DR SMART; SM00397; t-SNARE; 1.  
 DR PROSITE; P550192; T-SNARE; 1.  
 DR PROSITE; P550192; T-SNARE; 1.  
 SQ SEQUENCE 286 AA; 33028 MW; 5FB1D6B10B77DEAB CRC64;

Query Match 63.6%; Score 103; DB 2; Length 286;  
 Best Local Similarity 69.0%; Pred. No. 2.4e-05;  
 Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;





modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

CC EMBL: D10475; BAA01278.1; -

DR PIR: A38216; S51193.

DR HSBP; P32851; 1JTH.

DR WGD; MGI:108059; EpiM.

DR InterPro; IPR006012; Syntaxin.

DR InterPro; IPR006011; Syntaxin\_N.

DR InterPro; IPR010989; t-snare.

DR InterPro; IPR000727; T SNARE.

DR Pfam; PF05739; SNARE; 1.

DR Pfam; PF00804; Syntaxin; 1.

DR SMART; SM00503; SYN; 1.

DR SMART; SM00397; t-SNARE; 1.

DR PROSITE; PS00914; SYNTAXIN; 1.

DR PROSITE; PS0192; T SNARE; 1.

KW Coiled coil; Transmembrane.

FT DOMAIN 1 265 Cytoplasmic (Potential).

FT TRANSMEM 266 289 Anchor for type IV membrane protein (Potential).

FT DOMAIN 68 101 Coiled coil (Potential).

FT DOMAIN 192 254 t-SNARE coiled-coil homology.

SO SEQUENCE 289 AA; 33178 MW; 9D82330D0F5CA2F4 CRC64;

Query Match 63.6%; Score 103; DB 1; Length 289;  
Best Local Similarity 69.0%; Pred. No. 2.9e-05;  
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 TROALNEISARHSIGQLERSIRELHDF 30  
DB 188 TROALNEISARHSIGQLERSIRELHDF 216

RESULT 13

Q80W45 PRELIMINARY; PRT; 289 AA.

AC Q80W45;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-UTN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Epimorphin.

GN Name=Epim.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN NCBI [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Olfactory epithelium;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.K., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ussit T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McBean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman N., Madan A., Rodrigues S., Sanchez A.,

RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywiński M.I., Skalek U., Smalins D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]

RP SEQUENCE FROM N.A.

RC TISSUE=Olfactory epithelium;

RA Strausberg R.L.

RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

CC -1 SIMILARITY: Belongs to the syntaxin/epimorphin family.

DR EMBL: BC046279; AAH46279.1; -

DR HSBP; P32851; 1JTH.

DR WGD; MGI:108059; EpiM.

DR GO; GO:0016020; C:membrane, IEA.

DR GO; GO:0008565; F:protein transporter activity; IEA.

DR GO; GO:0006886; P:intracellular protein transport; IEA.

DR InterPro; IPR006012; Syntaxin\_N.

DR InterPro; IPR006011; Syntaxin.

DR InterPro; IPR010989; t-snare.

DR InterPro; IPR000727; T SNARE.

DR Pfam; PF05739; SNARE; 1.

DR Pfam; PF00804; Syntaxin; 1.

DR SMART; SM00503; SYN; 1.

DR SMART; SM00397; t-SNARE; 1.

DR PROSITE; PS00914; SYNTAXIN; 1.

DR PROSITE; PS0192; T SNARE; 1.

SO SEQUENCE 289 AA; 33177 MW; 77C29467AFBFCF8E CRC64;

Query Match 63.6%; Score 103; DB 2; Length 289;  
Best Local Similarity 69.0%; Pred. No. 2.9e-05;  
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 TROALNEISARHSIGQLERSIRELHDF 30  
DB 188 TROALNEISARHSIGQLERSIRELHDF 216

RESULT 14

RPMO RAT STANDARD; PRT; 290 AA.

AC P50279; Q08846; Q08847; Q08848;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Epimorphin (Syntaxin 2).

GN Name=Epim; Synonyms=Stx2;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN NCBI [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93386759; PubMed=7690687;

RA Bennett M.K., Garcia-Arreaga J.E., Elferink L.A., Peterson K.,

RA Fleming A.M., Hazuka C.D., Scheller R.H.,

RT "The syntaxin family of vesicular transport receptors."

RL Cell 74:863-873(1993).

RN NCBI [2]

RP SEQUENCE OF 11-290 FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Liver;

RX MEDLINE=97032890; PubMed=893452;

RA Zha H., Remmers B.F., Szpirer C., Szpirer J., Zhang H., Kozak C.A.,

RA Wilder R.L.;

RT "The epimorphin gene is highly conserved among humans, mice, and rats

RT and maps to human chromosome 7, mouse chromosome 5, and rat chromosome

RT 12."

RL Genomics 37:386-389(1996).

CC -1 FUNCTION: Essential for epithelial morphogenesis.

CC -1 SUBCELLULAR LOCATION: Type IV membrane protein.

CC -1 ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=3;

CC Comment=Additional isoforms seem to exist;

CC Name=2A;

CC IsoId=P50279-1; Sequence=Displayed;

CC Name=2B; Synonyms=2';

CC IsoId=P50279-2; Sequence=VSP\_006336;

CC Name=2C; Synonyms=2'';

CC IsoId=P50279-3; Sequence=VSP\_006337;

CC -1- TISSUE SPECIFICITY: Heart, spleen, liver, and testis.  
 CC -1- SIMILARITY: Belongs to the syntaxin/epimorphin family.  
 CC -1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L20823; AAA03044.1; -  
 DR EMBL: L20888; AAA03048.1; -  
 DR EMBL: L20888; AAA03048.1; -  
 DR EMBL: U35047; AAC52908.1; -  
 DR EMBL: U35039; AAC52908.1; JOINED.  
 DR EMBL: U35040; AAC52908.1; JOINED.  
 DR EMBL: U35041; AAC52908.1; JOINED.  
 DR EMBL: U35043; AAC52908.1; JOINED.  
 DR EMBL: U35044; AAC52908.1; JOINED.  
 DR EMBL: U35045; AAC52908.1; JOINED.  
 DR EMBL: U35046; AAC52908.1; JOINED.  
 DR HSSP: P32851; 1JTH.  
 DR RGD: 2558; Epim.  
 DR InterPro: IPR006012; Syntaxin.  
 DR InterPro: IPR006011; Syntaxin\_N.  
 DR InterPro: IPR010989; t-snare.  
 DR InterPro: IPR000727; t-SNARE.  
 DR Pfam: PF05739; SNARE; 1.  
 DR Pfam: PF08084; Syntaxin; 1.  
 DR SMART: SMO0503; SYN; 1.  
 DR SMART: SMO0397; t-SNARE; 1.  
 DR PROSITE: PS00914; SYNTAXIN; 1.  
 DR PROSITE: PS0192; t-SNARE; 1.  
 KM Alternative splicing; Coiled coil; Transmembrane.  
 FT DOMAIN 1 266  
 FT TRANSMEM 267 290  
 FT ANCHOR FOR TYPE IV MEMBRANE PROTEIN  
 FT (Potential).  
 FT COILED COIL (Potential).  
 FT t-SNARE coiled-coil homology.  
 FT KMWIAAVVAVIAVLAALIGLTGK -> KMFPLICVIT  
 FT LVIIGIILATLALS (in isoform 2B).  
 FT /FTID=VSP 006336.  
 FT KMWIAAVVAVIAVLAALIGLTGK -> GYLCALGRQC  
 FT (in isoform 2C).  
 FT /FTID=VSP 006337.  
 FT S -> R (in Ref. 2).  
 FT T -> S (in Ref. 2).  
 FT CONFLICT 146 146  
 FT CONFLICT 287 287  
 SQ SEQUENCE 290 AA; 33359 MW; 37FBAD7ADPD3841 CRC64;  
 Query Match 63.6%; Score 103; DB 1; Length 290;  
 Best Local Similarity 69.0%; Pred. No. 2.9e-05;  
 Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 TROALNEISARHSGIOQLERSIRELHDMF 30  
 DB 189 TROALNEISARHSGIOQLERSIRELHDMF 217  
 RESULT 15  
 OGDULT1  
 ID OGDULT1 PRELIMINARY; PRT; 141 AA.  
 AC OGDULT1;  
 DT 01-OCT-2004 (TREMBlrel. 28, Created)  
 DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE T-SNARE syntaxin 1B (Fragment).  
 OS Lateolabrax japonicus (Japanese sea perch) (Japanese sea bass).  
 OC Bakayocia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;  
 OC Moronidae; Lateolabrax.

OX NCBI TaxID=8164;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chen K., Huang X.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY63789; AAT67160.1; -  
 FT NON\_TER 1 1  
 FT NON\_TER 1 1  
 SQ SEQUENCE 141 AA; 16690 MW; 540C46F71E6204C CRC64;  
 Query Match 62.3%; Score 101; DB 2; Length 141;  
 Best Local Similarity 69.0%; Pred. No. 2.6e-05;  
 Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 TROALNEISARHSGIOQLERSIRELHDMF 30  
 DB 65 TROALNEISARHSGIOQLERSIRELHDMF 93

Search completed: November 10, 2004, 12:27:00  
 Job time : 38.1065 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 ; Search time 8.33951 Seconds  
(without alignments)  
222.664 Million cell updates/sec

Title: US-10-092-750-10  
Perfect score: 144  
Sequence: 1 MFSDIYGIREFADGLCEVGGKWSRPE 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 6631800 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A COMB.pap:\*  
2: /cgn2\_6/ptodata/1/1aa/6A COMB.pap:\*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pap:\*  
4: /cgn2\_6/ptodata/1/1aa/6B COMB.pap:\*  
5: /cgn2\_6/ptodata/1/1aa/6C COMB.pap:\*  
6: /cgn2\_6/ptodata/1/1aa/6D COMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	88.9	172	3	US-09-311-311C-27
2	60.5	42.0	167	2	US-08-690-849-2
3	60.5	42.0	167	3	US-09-004-053-2
4	49.5	34.4	191	4	US-09-270-767-43182
5	47	32.6	377	3	US-09-023-023-2
6	47	32.6	729	4	US-09-248-796A-17121
7	46	31.9	101	4	US-09-543-681A-4293
8	46	31.9	144	3	US-08-858-207A-523
9	46	31.9	168	4	US-09-737-300-2
10	46	31.9	186	4	US-09-252-991A-24198
11	46	31.9	1033	4	US-09-583-110-5097
12	46	31.9	1042	3	US-09-387-695-2
13	46	31.9	1651	4	US-09-543-681A-6604
14	45.5	31.6	100	4	US-09-107-532A-6916
15	45.5	31.6	155	3	US-08-821-278A-17
16	45	31.2	129	4	US-09-107-532A-5649
17	45	31.2	160	4	US-09-540-236-2514
18	45	31.2	301	4	US-09-134-000C-6225
19	45	31.2	531	2	US-08-975-114A-4
20	45	31.2	729	4	US-08-849-281A-4
21	45	31.2	731	4	US-09-107-532A-6946
22	44.5	30.9	776	4	US-09-134-000C-4364
23	44	30.6	112	4	US-09-796-766-2
24	44	30.6	114	4	US-09-270-767-32115
25	44	30.6	114	4	US-09-270-767-47332
26	44	30.6	233	4	US-09-248-796A-14148
27	44	30.6	252	4	US-09-796-766-4

28	44	30.6	759	2	US-08-637-759B-89	Sequence 89, Appl
29	44	30.6	759	3	US-08-871-355A-89	Sequence 89, Appl
30	44	30.6	759	4	US-09-201-945-89	Sequence 89, Appl
31	43.5	30.2	422	3	US-09-270-767-42572	Sequence 42572, A
32	43	29.9	220	4	US-09-489-039A-7578	Sequence 7578, Ap
33	43	29.9	277	4	US-09-602-787A-266	Sequence 266, Appl
34	43	29.9	358	4	US-09-489-039A-7189	Sequence 7189, Ap
35	43	29.9	427	4	US-09-252-991A-30420	Sequence 30420, A
36	43	29.9	532	4	US-09-543-681A-7614	Sequence 7614, Appl
37	43	29.9	959	4	US-09-248-796A-20776	Sequence 20776, A
38	42.5	29.5	122	4	US-09-328-352-5206	Sequence 5206, Ap
39	42.5	29.5	425	4	US-09-538-092-1366	Sequence 1366, Ap
40	42.5	29.5	460	4	US-09-543-681A-7248	Sequence 7248, Ap
41	42.5	29.5	524	4	US-09-408-020-10	Sequence 10, Appl
42	42	29.2	64	4	US-09-489-039A-7726	Sequence 7726, Ap
43	42	29.2	140	4	US-09-248-796A-18372	Sequence 18372, A
44	42	29.2	180	4	US-09-270-767-45651	Sequence 45651, A
45	42	29.2	300	4	US-09-252-991A-28063	Sequence 28063, A

# ALIGNMENTS

RESULT 1	US-09-311-311C-27	US-09-311-311C-27
Sequence 27, Appl	Application US/09311311C	
Patent No. 6358738		
GENERAL INFORMATION:		
APPLICANT: Erikson, et al.		
TITLE OF INVENTION: POLO BOX THERAPEUTIC COMPOSITIONS,		
FILE REFERENCE: 1874/117		
CURRENT APPLICATION NUMBER: US/09/311, 311C		
CURRENT FILING DATE: 1999-05-13		
PRIOR APPLICATION NUMBER: US 60/085,296		
PRIOR FILING DATE: 1998-05-13		
NUMBER OF SEQ ID NOS: 27		
SOFTWARE: FastSeq for Windows Version 4.0		
SEQ ID NO 27		
LENGTH: 172		
TYPE: PRT		
ORGANISM: Homo sapiens		
FEATURE:		
NAME/KEY: DOMAIN		
LOCATION: (1)...		
OTHER INFORMATION: TCTP/23 protein		
US-09-311-311C-27		
Query Match	88.9%	Score 128, DB 3, Length 172,
Best Local Similarity	92.9%	Pred. No. 2.7e-13,
Matches 26, Conservative	0, Mismatches	2, Indels 0, Gaps 0;
Db	1 MFSDIYGIREFADGLCEVGGKWSRPE 28	
	13 MFSDIYGIREFADGLCEVGGKWSRTE 40	
RESULT 2	US-08-690-849-2	
Sequence 2, Appl	Application US/08690849	
Patent No. 5952194		
GENERAL INFORMATION:		
APPLICANT: Stiegler, Gary L.		
TITLE OF INVENTION: ECTOPARASITE HISTAMINE RELEASING		
FILE REFERENCE: FACTOR,		
TITLE OF INVENTION: GENES AND USES THEREOF		
NUMBER OF SEQUENCES: 5		
CORRESPONDENCE ADDRESS:		
ADDRESS: Sheridan Ross & McIntosh,		
STREET: 1700 Lincoln Street, Suite 3500		
CITY: Denver		
STATE: CO		
COUNTRY: U.S.A.		

```

; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,849
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-49
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-9700
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-690-849-2

Query Match          42.0%; Score 60.5; DB 2; Length 167;
Best Local Similarity 50.0%; Pred. No. 0.038;
Matches 14; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 1 MFSDYIGIREIADGCLFVEGKRVSRPE 28
DB 8 MFSDTYKIK-LVDEVLVEYVTKLVSRSQ 34

RESULT 3
US-09-004-053-2
; Sequence 2, Application US/09004053
; Patent No. 6063902
; GENERAL INFORMATION:
; APPLICANT: Stiegler, Gary L.
; TITLE OF INVENTION: ECTOPARASITE HISTAMINE RELEASING
; TITLE OF INVENTION: FACTOR.
; TITLE OF INVENTION: GENES AND USES THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,053
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/690,849
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-004-053-2

Query Match          42.0%; Score 60.5; DB 3; Length 167;
Best Local Similarity 50.0%; Pred. No. 0.038;
Matches 14; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 1 MFSDYIGIREIADGCLFVEGKRVSRPE 28
DB 8 MFSDTYKIK-LVDEVLVEYVTKLVSRSQ 34

RESULT 4
US-09-270-767-43182
; Sequence 43182, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 43182
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-09-270-767-43182

Query Match          34.4%; Score 49.5; DB 4; Length 191;
Best Local Similarity 38.5%; Pred. No. 3;
Matches 10; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

QY 1 MFSDYIGIREIADGCLFVEGKRVSR 26
DB 32 MFADTYKIK-LVDDVVEYVYKULTR 56

RESULT 5
US-09-023-023-2
; Sequence 2, Application US/09023023
; Patent No. 6121018
; GENERAL INFORMATION:
; APPLICANT: Kristine Kay Kikly
; TITLE OF INVENTION: Interleukin-1 Beta Converting Enzyme Like Apoptosis Pr
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,023
; FILING DATE: 12-FEB-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/042,030
; FILING DATE: March 27, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50013
; TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 610-270-5219  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-023-023-2

Query Match 32.6%; Score 47; DB 3; Length 377;  
Best Local Similarity 40.0%; Pred. No. 17;  
Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1;

Qy 4 DIVGIRIADGL--CLEVEGKAVSR 26  
Db 158 DILGKMLLEGLGTYVEVEKLTAR 182

RESULT 6  
US-09-248-796A-17121  
Sequence 17121, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstein et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 17121  
LENGTH: 729  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-17121

Query Match 32.6%; Score 47; DB 4; Length 729;  
Best Local Similarity 62.5%; Pred. No. 38;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 FSDIYGIREDGLCL 17  
Db 85 FSNLYGIPEHADSLH 100

RESULT 7  
US-09-543-681A-4293  
Sequence 4293, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 4293  
LENGTH: 101  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-4293

Query Match 31.9%; Score 46; DB 4; Length 101;  
Best Local Similarity 56.0%; Pred. No. 5.3;

Matches 14; Conservative 4; Mismatches 5; Indels 2; Gaps 2;  
Qy 3 SDIY-GIRIADGLCL-VEGKAVS 25  
Db 53 SDIYGRDIAEQLNAGAVNGKIVS 77

RESULT 8  
US-08-858-207A-523  
Sequence 523, Application US/08858207A  
Patent No. 6348328  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
APPLICANT: Hodgson, John  
APPLICANT: Knowles, David  
APPLICANT: Nicholas, Richard  
TITLE OF INVENTION: No. 6348328el Compounds  
NUMBER OF SEQUENCES: 552  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/858,207A  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/017670  
FILING DATE: 14-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmi, Edward R.  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50475  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 523:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 144 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6348328e  
US-08-858-207A-523

Query Match 31.9%; Score 46; DB 3; Length 144;  
Best Local Similarity 42.9%; Pred. No. 8.1;  
Matches 12; Conservative 6; Mismatches 8; Indels 2; Gaps 1;

Qy 1 FSDIYGI--RIADGLCLEVEGKAVSR 26  
Db 34 LFSDLYRQVGOEIKEGAFYVKKIQSR 61

RESULT 9  
US-09-737-300-2  
Sequence 2, Application US/09737300  
Patent No. 6545902  
GENERAL INFORMATION:  
APPLICANT: KOREA KUMHO PETROCHEMICAL CO., LTD.  
TITLE OF INVENTION: Transgenic plants and plant cells with improved growth rate and  
FILE REFERENCE: px00121/us

CURRENT APPLICATION NUMBER: US/09/737,300  
CURRENT FILING DATE: 2000-12-15  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Kopatentln 1.71  
SEQ ID NO 2  
LENGTH: 168  
TYPE: PRT  
ORGANISM: nttcpe protein(tabacco translationally controlled tumor protein)  
US-09-737-300-2

Query Match 31.9%; Score 46; DB 4; Length 168;  
Best Local Similarity 40.9%; Pred. No. 9.7;  
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 SDIYGIETADGCLCEVEGKWSR 24  
DB 15 SDSFSTLENGVMEVQKRW 36

RESULT 10  
US-09-252-991A-24198  
Sequence 24198, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 24198  
LENGTH: 186  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24198

Query Match 31.9%; Score 46; DB 4; Length 186;  
Best Local Similarity 33.3%; Pred. No. 11;  
Matches 8; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 5 IYGIETADGCLCEVEGKWSRPE 28  
DB 146 MYGLQKQSDGMSVSGSSYSDE 169

RESULT 11  
US-09-583-110-5097  
Sequence 5097, Application US/09583110  
Patent No. 6699703  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al.  
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
FILE REFERENCE: P4TH00-07A  
CURRENT APPLICATION NUMBER: US/09/583,110  
CURRENT FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/107,433  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/085,131  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: US 60/051,553  
PRIOR FILING DATE: 1997-07-02  
NUMBER OF SEQ ID NOS: 5322  
SEQ ID NO 5097  
LENGTH: 1033  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-583-110-5097

Query Match 31.9%; Score 46; DB 4; Length 1033;  
Best Local Similarity 42.9%; Pred. No. 84;  
Matches 12; Conservative 6; Mismatches 8; Indels 2; Gaps 1;

QY 1 MFSDIYGI--REIADGCLCEVEGKWSR 26  
DB 923 LFSDIYRQVQGEIKGAFYVVKGIQSR 950

RESULT 12  
US-09-387-695-2  
Sequence 2, Application US/09387695  
Patent No. 6280990  
GENERAL INFORMATION:  
APPLICANT: May, Earl  
APPLICANT: Van Horn, Stephanie  
APPLICANT: Warren, Patrick V.  
APPLICANT: Warren, Richard L.  
TITLE OF INVENTION: dnaB  
FILE REFERENCE: GMI0237  
CURRENT APPLICATION NUMBER: US/09/387,695  
CURRENT FILING DATE: 1999-08-31  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 1042  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-387-695-2

Query Match 31.9%; Score 46; DB 3; Length 1042;  
Best Local Similarity 42.9%; Pred. No. 84;  
Matches 12; Conservative 6; Mismatches 8; Indels 2; Gaps 1;

QY 1 MFSDIYGI--REIADGCLCEVEGKWSR 26  
DB 932 LFSDIYRQVQGEIKGAFYVVKGIQSR 959

RESULT 13  
US-09-543-681A-6604  
Sequence 6604, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 6604  
LENGTH: 1651  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-6604

Query Match 31.9%; Score 46; DB 4; Length 1651;  
Best Local Similarity 40.7%; Pred. No. 1.5e+02;  
Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 2 FSDIYGIETADGCLCEVEGKWSRPE 28  
DB 943 YSIIYAASKRQDGLTEILKALSVPE 969

RESULT 14  
US-09-107-532A-6916  
Sequence 6916, Application US/09107532A  
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Ducelette-Stamm and David Bush  
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO  
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 7310  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: GENOME THERAPEUTICS CORPORATION  
 STREET: 100 Beaver Street  
 CITY: Waltham  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02354  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD-ROM ISO9660  
 COMPUTER: PC  
 OPERATING SYSTEM: <Unknown>  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/107,532A  
 FILING DATE: 30-Jun-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/085,598  
 FILING DATE: 14 May 1998  
 APPLICATION NUMBER: 60/051571  
 FILING DATE: July 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ariniello, Pamela Deneke  
 REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: GTC-012  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781)893-5007  
 TELEFAX: (781)893-8277  
 INFORMATION FOR SEQ ID NO: 6916:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 100 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEetical: YES  
 ORIGINAL SOURCE:  
 ORGANISM: Enterococcus faecium  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (B) LOCATION 1..100  
 SEQUENCE DESCRIPTION: SEQ ID NO: 6916:  
 US-09-107-532A-6916

Query Match 31.6% Score 45.5; DB 4; Length 100;  
 Best Local Similarity 37.0%; Pred. No. 6.3; 10; Indels 3; Gaps 1;  
 Matches 10; Conservative 4; Mismatches 10; Indels 3; Gaps 1;  
 QY 2 FSDIYGRIRI--ADGICLVEYEGKMSRPE 28  
 DB 31 FGKVGKTEVHDGFSI--GMSVEQPE 54

RESULT 15  
 US-08-821-278A-17  
 Sequence 17, Application US/08821278A  
 Patent No. 6238902  
 GENERAL INFORMATION:  
 APPLICANT: Cheng, Jili  
 APPLICANT: Lasky, Laurence A.  
 TITLE OF INVENTION: Protein Tyrosine Phosphatases  
 FILE REFERENCE: P1010R1  
 CURRENT APPLICATION NUMBER: US/08/821,278A  
 CURRENT FILING DATE: 1997-03-20  
 NUMBER OF SEQ ID NOS: 23  
 SEQ ID NO 17  
 LENGTH: 155  
 TYPE: prt  
 ORGANISM: Homo Sapien  
 US-08-821-278A-17

Query Match 31.6% Score 45.5; DB 3; Length 155;  
 Best Local Similarity 48.3%; Pred. No. 11;  
 Matches 14; Conservative 1; Mismatches 9; Indels 5; Gaps 2;

QY 2 FSDIYGRIRI--ADGICLVEYEGKMSRPE 28  
 DB 14 FSDIQASAAKADGVCSIVAG--SRPE 39

Search completed: November 10, 2004, 12:32:24  
 Job time: 9.33951 secs

[This page blank (uspio)]



Fri Nov 12 14:54:45 2004

us-10-092-750-10.rapb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 ; Search time 27.0662 Seconds  
(without alignments)  
366.225 Million cell updates/sec

Title: US-10-092-750-10  
Perfect score: 144  
Sequence: 1 MFSDIVGIREIADGLCLEVEGKRVSRPE 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225866 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PC1\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PC1US\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	144	100.0	28	US-10-092-750-10	Sequence 10, Appl
2	128	88.9	78	US-10-425-115-239469	Sequence 239469,
3	128	88.9	86	US-10-264-049-3470	Sequence 3470, Ap
4	128	88.9	172	US-10-153-668-350	Sequence 350, App
5	128	88.9	172	US-10-021-753-2	Sequence 2, Appl1
6	128	88.9	172	US-10-139-794-6	Sequence 6, Appl1
7	125	86.8	172	US-10-021-753-4	Sequence 4, Appl1
8	125	86.8	172	US-10-451-861-19	Sequence 19, Appl
9	125	86.8	172	US-10-451-861-20	Sequence 20, Appl
10	121	84.0	172	US-10-021-753-3	Sequence 3, Appl1
11	119	82.6	172	US-10-021-753-5	Sequence 5, Appl1
12	90	62.5	54	US-10-425-115-308966	Sequence 308966,
13	85	59.0	61	US-10-425-115-336562	Sequence 336562,

14	62	43.1	79	US-10-424-599-215480	Sequence 215480,
15	57	39.6	77	US-10-425-115-276150	Sequence 276150,
16	57	39.6	167	US-10-767-701-47116	Sequence 47116, A
17	57	39.6	167	US-10-425-115-314609	Sequence 314609, A
18	57	39.6	180	US-10-425-114-66045	Sequence 66045, A
19	55	38.2	96	US-10-424-599-197485	Sequence 197485,
20	55	38.2	168	US-10-437-963-204381	Sequence 204381,
21	54.5	37.8	172	US-10-291-172-317	Sequence 317, App
22	54.5	37.8	172	US-10-291-172-317	Sequence 317, App
23	54.5	37.8	172	US-10-291-172-317	Sequence 317, App
24	54.5	37.8	178	US-10-221-278-693	Sequence 693, App
25	54	37.5	102	US-10-424-599-215024	Sequence 215024,
26	54	37.5	103	US-10-424-599-215533	Sequence 215533,
27	54	37.5	107	US-09-864-408A-7036	Sequence 2036, Ap
28	54	37.5	114	US-10-424-599-269163	Sequence 269163,
29	54	37.5	115	US-10-424-599-269163	Sequence 269163,
30	54	37.5	130	US-10-424-599-269165	Sequence 269165,
31	54	37.5	144	US-10-424-599-204067	Sequence 204067,
32	54	37.5	160	US-10-424-599-213300	Sequence 213300,
33	54	37.5	168	US-10-021-753-9	Sequence 9, Appl1
34	54	37.5	168	US-10-424-599-269167	Sequence 269167,
35	54	37.5	168	US-10-424-599-269172	Sequence 269172,
36	54	37.5	168	US-10-437-963-115017	Sequence 115017,
37	54	37.5	793	US-10-369-493-10366	Sequence 10366, A
38	53.5	37.2	141	US-10-424-599-157526	Sequence 157526,
39	53.5	37.2	167	US-10-425-115-229477	Sequence 229477,
40	53	36.8	59	US-10-424-599-211800	Sequence 211800,
41	53	36.8	64	US-10-425-115-195430	Sequence 195430,
42	53	36.8	78	US-10-424-599-148160	Sequence 148160,
43	53	36.8	81	US-10-425-115-264849	Sequence 264849,
44	53	36.8	111	US-10-425-114-60200	Sequence 60200, A
45	53	36.8	128	US-10-425-114-61941	Sequence 61941, A

#### ALIGNMENTS

RESULT 1  
US-10-092-750-10  
; Sequence 10, Application US/10092750  
; Publication No. US20030032157A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; APPLICANT: Wright, Martin C.  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL  
; FILE REFERENCE: 50036/050002  
; CURRENT APPLICATION NUMBER: US/10/092, 750  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: US 60/274, 526  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-750-10

Query Match 100.0%; Score 144; DB 14;  
Best Local Similarity 100.0%; Pred. No. 5.8e-15;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 1 MFSDIVGIREIADGLCLEVEGKRVSRPE 28  
Db 1 MFSDIVGIREIADGLCLEVEGKRVSRPE 28  
RESULT 2  
US-10-425-115-239469  
; Sequence 239469, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:



```

; Sequence 2, Application US/10021753
; Publication No. US20030172388A1
; GENERAL INFORMATION:
; APPLICANT: FUJISE, KEN
; APPLICANT: YEH, EDWARD T.H.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO FORTILIN, AN
; TITLE OF INVENTION: ANTI-APOPTOTIC MOLECULE, AND MODULATORS OF FORTILIN
; FILE REFERENCE: UTS#251US
; CURRENT APPLICATION NUMBER: US/10/021,753
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-753-2

```

```

Query Match      88.9%; Score 128; DB 14; Length 172;
Best Local Similarity 92.9%; Pred. No. 1.4e-11;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 MFSDIYKIREIADGCLCEVEGKMSRPE 28
Db 13 MFSDIYKIREIADGCLCEVEGKMSRTE 40

```

```

RESULT 6
US-10-139-794-6
; Sequence 6, Application US/10139794
; Publication No. US20030232421A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS, LYNX THERAPEUTICS INC.
; APPLICANT: Pierre Legrain, Simon Whiteside, Jen-I Mao, Irina Khrebukova, Shujun Luc
; TITLE OF INVENTION: Protein-Protein Interactions In Adipocyte Cells (3)
; FILE REFERENCE: B4883A
; CURRENT APPLICATION NUMBER: US/10/139,794
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US 60/288,885
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 2930
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 6
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Translation of SEQ ID NO:5
US-10-139-794-6

```

```

Query Match      88.9%; Score 128; DB 14; Length 172;
Best Local Similarity 92.9%; Pred. No. 1.4e-11;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 MFSDIYKIREIADGCLCEVEGKMSRPE 28
Db 13 MFSDIYKIREIADGCLCEVEGKMSRTE 40

```

```

RESULT 7
US-10-021-753-4
; Sequence 4, Application US/10021753
; Publication No. US20030172388A1
; GENERAL INFORMATION:
; APPLICANT: FUJISE, KEN
; APPLICANT: YEH, EDWARD T.H.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO FORTILIN, AN
; TITLE OF INVENTION: ANTI-APOPTOTIC MOLECULE, AND MODULATORS OF FORTILIN
; FILE REFERENCE: UTS#251US
; CURRENT APPLICATION NUMBER: US/10/021,753
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1

```

```

; SEQ ID NO 4
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-021-753-4

```

```

Query Match      86.8%; Score 125; DB 14; Length 172;
Best Local Similarity 89.3%; Pred. No. 4.1e-11;
Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 MFSDIYKIREIADGCLCEVEGKMSRPE 28
Db 13 MFSDIYKIREIADGCLCEVEGKMSRTE 40

```

```

RESULT 8
US-10-451-861-19
; Sequence 19, Application US/10451861
; Publication No. US20040175772A1
; GENERAL INFORMATION:
; APPLICANT: AMSON, Robert
; APPLICANT: TELERMAN, Adam
; APPLICANT: PASSER, Brent
; TITLE OF INVENTION: SCREENING METHOD BASED ON TSAP6 BINDING PARTNERS
; FILE REFERENCE: 11416-011-999 (344 910-US)
; CURRENT APPLICATION NUMBER: US/10/451,861
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: PCT/FR 01/04 188
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR 01/02 896
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: FR 00/17 027
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 19
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: Mouse TCTP NP_033455
US-10-451-861-19

```

```

Query Match      86.8%; Score 125; DB 16; Length 172;
Best Local Similarity 89.3%; Pred. No. 4.1e-11;
Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 MFSDIYKIREIADGCLCEVEGKMSRPE 28
Db 13 MFSDIYKIREIADGCLCEVEGKMSRTE 40

```

```

RESULT 9
US-10-451-861-20
; Sequence 20, Application US/10451861
; Publication No. US20040175772A1
; GENERAL INFORMATION:
; APPLICANT: AMSON, Robert
; APPLICANT: TELERMAN, Adam
; APPLICANT: PASSER, Brent
; TITLE OF INVENTION: SCREENING METHOD BASED ON TSAP6 BINDING PARTNERS
; FILE REFERENCE: 11416-011-999 (344 910-US)
; CURRENT APPLICATION NUMBER: US/10/451,861
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: PCT/FR 01/04 188
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR 01/02 896
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: FR 00/17 027
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 20

```

LENGTH: 172  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: Human TCTP- NP\_003286  
 US-10-451-861-20

Query Match 86.8%; Score 125; DB 16; Length 172;  
 Best Local Similarity 89.3%; Pred. No. 4.1e-11;  
 Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MFSDYIGREIADGCLCEVEGKMSRPE 28  
 |||||  
 Db 13 LFSDIKIREIADGCLCEVEGKMSRTE 40

RESULT 10  
 US-10-021-753-3  
 Sequence 3, Application US/10021753  
 Publication No. US20030172388A1  
 GENERAL INFORMATION:  
 APPLICANT: FUJISE, KEN  
 APPLICANT: YEH, EDWARD T.H.  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO FORTILIN, AN  
 TITLE OF INVENTION: ANTI-APOPTOTIC MOLECULE, AND MODULATORS OF FORTILIN  
 FILE REFERENCE: UTSH:251US  
 CURRENT APPLICATION NUMBER: US/10/021,753  
 CURRENT FILING DATE: 2001-10-30  
 NUMBER OF SEQ ID NOS: 9  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 3  
 LENGTH: 172  
 TYPE: PRT  
 ORGANISM: Rabbit  
 US-10-021-753-3

Query Match 84.0%; Score 121; DB 14; Length 172;  
 Best Local Similarity 89.3%; Pred. No. 1.7e-10;  
 Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MFSDYIGREIADGCLCEVEGKMSRPE 28  
 |||||  
 Db 13 MFSDYKIREIADGCLCEVEGKMSRTE 40

RESULT 11  
 US-10-021-753-5  
 Sequence 5, Application US/10021753  
 Publication No. US20030172388A1  
 GENERAL INFORMATION:  
 APPLICANT: FUJISE, KEN  
 APPLICANT: YEH, EDWARD T.H.  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO FORTILIN, AN  
 TITLE OF INVENTION: ANTI-APOPTOTIC MOLECULE, AND MODULATORS OF FORTILIN  
 FILE REFERENCE: UTSH:251US  
 CURRENT APPLICATION NUMBER: US/10/021,753  
 CURRENT FILING DATE: 2001-10-30  
 NUMBER OF SEQ ID NOS: 9  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 5  
 LENGTH: 172  
 TYPE: PRT  
 ORGANISM: Chicken  
 US-10-021-753-5

Query Match 82.6%; Score 119; DB 14; Length 172;  
 Best Local Similarity 82.1%; Pred. No. 3.5e-10;  
 Matches 23; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MFSDYIGREIADGCLCEVEGKMSRPE 28  
 |||||  
 Db 13 MFSDYKIREIADGCLCEVEGKMSRTE 40

RESULT 12  
 US-10-425-115-308966  
 Sequence 308966, Application US/10425115  
 Publication No. US20040214272A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 TITLE OF INVENTION: Plants  
 FILE REFERENCE: 38-21(53222)B  
 CURRENT APPLICATION NUMBER: US/10/425,115  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO 308966  
 LENGTH: 54  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (1)..(54)  
 OTHER INFORMATION: unsure at all Xaa locations  
 FEATURE:  
 OTHER INFORMATION: Clone ID: MRT4577\_44843C.1.pep  
 US-10-425-115-308966

Query Match 62.5%; Score 90; DB 17; Length 54;  
 Best Local Similarity 94.7%; Pred. No. 2.9e-06;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 IREIADGCLCEVEGKMSR 26  
 |||||  
 Db 2 IREIADGCLCEVEGKMSXR 20

RESULT 13  
 US-10-425-115-336562  
 Sequence 336562, Application US/10425115  
 Publication No. US20040214272A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 TITLE OF INVENTION: Plants  
 FILE REFERENCE: 38-21(53222)B  
 CURRENT APPLICATION NUMBER: US/10/425,115  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO 336562  
 LENGTH: 61  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (1)..(61)  
 OTHER INFORMATION: unsure at all Xaa locations  
 FEATURE:  
 OTHER INFORMATION: Clone ID: MRT4577\_70059C.1.pep  
 US-10-425-115-336562

Query Match 59.0%; Score 85; DB 17; Length 61;  
 Best Local Similarity 74.1%; Pred. No. 2e-05;  
 Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 FSDYIGREIADGCLCEVEGKMSRPE 28  
 |||||  
 Db 21 FSDYKIREIADGCLCEVEGKMSRTE 47

## RESULT 14

US-10-424-599-215480  
; Sequence 215480, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 215480  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(79)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_36605C.1.pep  
US-10-424-599-215480

## Query Match

43.1%; Score 62; DB 15; Length 79;  
Best Local Similarity 50.0%; Pred. No. 0.036;

Matches 13; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 MFSDYGIREFIADGICLFEVGRMYSR 26

Db 12 LFSDFCCREIENGFLMEVGRKVGK 37

## RESULT 15

US-10-425-115-276150  
; Sequence 276150, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 276150  
; LENGTH: 77  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_183434C.1.pep  
US-10-425-115-276150

## Query Match

39.6%; Score 57; DB 17; Length 77;  
Best Local Similarity 44.0%; Pred. No. 0.56;

Matches 11; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

OY 1 MFSDYGIREFIADGICLFEVGRMYS 25

Db 13 LLSDFTYKELANGVLMEVGRKVT 37

Search completed: November 10, 2004, 16:35:46  
Job time : 28.0062 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:36:51; Search time 5.66049 Seconds  
(without alignments)  
475.942 Million cell updates/sec

Title: US-10-092-750-10

Perfect score: 144  
Sequence: 1 MFSDIYGIKREIADGLCLEVEGKMSRPE 28

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

1: p1r1.\*  
2: p1r2.\*  
3: p1r3.\*  
4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	88.9	172	1	S06590
2	125	86.8	172	1	S00775
3	119	82.6	172	1	A38950
4	116	80.6	172	1	A38956
5	61	42.4	157	2	S22489
6	61	42.4	157	2	S22489
7	58	40.3	167	2	T09686
8	54	37.5	168	1	A38958
9	54	37.5	299	2	S34588
10	50	34.7	299	1	S60035
11	50	34.7	299	1	S72173
12	50	34.7	435	2	E70711
13	49.5	34.4	235	2	F84054
14	45	34.0	197	2	A64484
15	49	34.0	366	2	B72725
16	48	33.3	542	2	A70826
17	47.5	33.0	77	2	A72256
18	47.5	33.0	421	2	AB2959
19	47	32.6	309	1	A69073
20	47	32.6	338	2	G69061
21	47	32.6	381	2	T10707
22	47	32.6	682	2	AE0033
23	47	32.6	816	2	T51813
24	47	32.6	816	2	B84605
25	47	32.6	3351	2	T13812
26	46.5	32.3	376	2	AD2585
27	46.5	32.3	376	2	C97367
28	46.5	32.3	388	2	AG3256
29	46.5	32.3	637	2	B82175

30	46	31.9	168	1	A38959	IGF-dependent hist
31	46	31.9	230	2	D81223	conserved hypothet
32	46	31.9	230	2	C81994	hypothetical prote
33	46	31.9	233	2	B71241	hypothetical prote
34	46	31.9	275	2	E69377	ABC transporter, A
35	46	31.9	277	2	C82128	4-aminu-4-deoxycho
36	46	31.9	361	2	G87706	hypothetical prote
37	46	31.9	412	2	G64059	probable serine tr
38	46	31.9	415	2	AG2928	glucarate dehydrat
39	46	31.9	417	2	AB3094	cell division prot
40	46	31.9	422	2	G98353	probable glucarate
41	46	31.9	1042	2	B95103	DNA polymerase III
42	46	31.9	1042	2	C97971	DNA-directed DNA p
43	45.5	31.6	331	2	AC1191	lipate protein 11
44	45.5	31.6	331	2	AC1549	lipate protein 11
45	45.5	31.6	385	2	T18712	hypothetical prote

## ALIGNMENTS

RESULT 1  
S06590  
IGF-dependent histamine-releasing factor - human  
N:Alternate names: 21k tumor protein; tumor-associated protein  
C/Species: Homo sapiens (man)  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: S06590; A43082; F33178  
R/Gross, B.; Gaestel, M.; Boehm, H.; Bielik, H.  
Nucleic Acids Res. 17, 8367, 1989  
A/Title: CDNA sequence coding for a translationally controlled human tumor protein.  
A/Reference number: S06590; M01D:90045593; PMID:2813067  
A/Accession: S06590  
A/Molecule type: mRNA  
A/Residues: 1-172 <GRO>  
A/Cross-references: UNIPROT:P13693; EMBL:X16064; NID:G37495; PIDD:CAA34200.1; P1D:G37496  
R/MacDonald, S.M.; Rafnar, T.; Langdon, J.; Lichtenstein, L.M.  
Science 269, 688-690, 1995  
A/Title: Molecular identification of an IGF-dependent histamine-releasing factor.  
A/Reference number: A43082; M01D:95350660; PMID:7542803  
A/Accession: A43082  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-18 <MAC>  
R/Ward, L.D.; Hong, J.; Whitehead, R.H.; Simpson, R.J.  
Electrophoresis 11, 883-891, 1990  
A/Title: Development of a database of amino acid sequences for human colon carcinoma prot  
C/Suprafamily: IGF-dependent histamine-releasing factor  
C/Keywords: lymphocyte  
A/Accession: F33178  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-18 <MAR>  
C/Genetics:  
A/Genes: GDB:RPT1; TCTP  
A/Cross-references: GDB:134697  
A/Map position: 13q14.3-13q14.3  
C/Suprafamily: IGF-dependent histamine-releasing factor  
Query Match: 88.9%; Score 128; DB 1; Length 172;  
Best Local Similarity: 92.9%; Pred. No. 8.7e-12;  
Matches: 26; Conservative: 0; Mismatches: 2; Indels: 0; Gaps: 0;

QY 1 MFSDIYGIKREIADGLCLEVEGKMSRPE 28  
DB 13 MFSDIYKIREIADGLCLEVEGKMSRTE 40

RESULT 2  
S00775  
IGF-dependent histamine-releasing factor - mouse  
N:Alternate names: 21k tumor protein; tumor-associated protein  
C/Species: Mus musculus (house mouse)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: S00775  
 R/Chitpatima, S.T.; Makrides, S.; Bandypadhyay, R.; Braverman, G.  
 Nucleic Acids Res. 16, 2350, 1988  
 A/Title: Nucleotide sequence of a major messenger RNA for a 21 kilodalton polypeptide h  
 A/Reference number: S00775; MUID:88189840; PMID:3357792  
 A/Accession: S00775  
 A/Molecule type: mRNA  
 A/Residues: 1-172 <CH>  
 A/Cross-references: UNIPROT:P14701, EMBL:X06407, NID:952848, PIND:CAA9697.1, PID:952848  
 C/Superfamily: IGE-dependent histamine-releasing factor  
 C/Keywords: lymphocyte

Query Match 86.8%; Score 125; DB 1; Length 172;  
 Best Local Similarity 89.3%; Pred. No. 2.5e-11;  
 Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 13 LFSDIYKIREIADGICLVEGKMWSRPE 40

RESULT 3  
 A38960  
 IGE-dependent histamine-releasing factor homolog - chicken  
 N/Alternate names: 21k tumor protein homolog; translationally-controlled tumor protein h  
 C/Species: Gallus gallus (chicken)  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: A38960  
 R/Sawada, K.; Agata, K.; Eguchi, G.  
 Submitted to GenBank, December 1993  
 A/Description: Analysis of the cDNA library of chicken lens fibers: identification of a  
 A/Reference number: A38960  
 A/Accession: A38960  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-172 <NIS>  
 A/Cross-references: UNIPROT:P43347, GB:D86312, NID:9517090, PIND:BA05374.1, PID:9517090  
 C/Superfamily: IGE-dependent histamine-releasing factor

Query Match 82.6%; Score 119; DB 1; Length 172;  
 Best Local Similarity 82.1%; Pred. No. 2e-10;  
 Matches 23; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 13 MFSDIYKIREIADGICLVEGKMWTRTE 40

RESULT 4  
 A38956  
 IGE-dependent histamine-releasing factor - rabbit  
 N/Alternate names: 21k tumor protein; tumor-associated protein  
 C/Species: Oryctolagus cuniculus (domestic rabbit)  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: A38956  
 R/Dawson, S.P.  
 Submitted to GenBank, November 1994  
 A/Description: A rabbit gene encoding a protein homologous to a translationally control  
 A/Reference number: A38956  
 A/Accession: A38956  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-172 <DAW>  
 A/Cross-references: UNIPROT:P43348, EMBL:Z46805, NID:9539941, PIND:CAA8626.1, PID:953994  
 C/Superfamily: IGE-dependent histamine-releasing factor  
 C/Keywords: lymphocyte

Query Match 80.6%; Score 116; DB 1; Length 172;  
 Best Local Similarity 88.9%; Pred. No. 5.7e-10;  
 Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 2 FSDIYKIREIADGICLVEGKMWSRPE 28

Db 14 LFSDIYKIREIADGICLVEGKMWSRTE 40

RESULT 5  
 S22489  
 IGE-dependent histamine-releasing factor homolog - alfalfa (fragment)  
 N/Alternate names: 21k tumor protein homolog; translationally-controlled tumor protein h  
 C/Species: Medicago sativa (alfalfa)  
 C/Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
 C/Accession: S22489; S19895  
 R/Pay, A.; Heberle-Bors, E.; Hirt, H.  
 Plant Mol. Biol. 19, 501-503, 1992  
 A/Title: An alfalfa cDNA encodes a protein with homology to translationally controlled h  
 A/Reference number: S22489; MUID:92322983; PMID:1623194  
 A/Accession: S22489  
 A/Molecule type: mRNA  
 A/Residues: 1-157 <PAY>  
 A/Cross-references: UNIPROT:P28014, EMBL:X63872, NID:919657, PIND:CAA4549.1, PID:919658  
 C/Superfamily: IGE-dependent histamine-releasing factor

Query Match 42.4%; Score 61; DB 2; Length 157;  
 Best Local Similarity 46.2%; Pred. No. 0.11;  
 Matches 12; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Db 3 LLSDSYKREIENGMLWEVGKMWTK 28

RESULT 6  
 T09686  
 TCP protein homolog - alfalfa  
 C/Species: Medicago sativa (alfalfa)  
 C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C/Accession: T09686  
 R/Sececi, J.; Nemeth, K.; Dudits, D.; Gyorgyey, J.  
 Submitted to the EMBL Data Library, June 1996  
 A/Description: Alfalfa homologue of TCP exhibits constitutive and abundant expression in  
 A/Reference number: Z16821  
 A/Accession: T09686  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-167 <SCB>  
 A/Cross-references: UNIPROT:P28014, EMBL:X98618  
 A/Experimental source: sub species Regen S, strain RA3  
 C/Superfamily: IGE-dependent histamine-releasing factor

Query Match 42.4%; Score 61; DB 2; Length 167;  
 Best Local Similarity 46.2%; Pred. No. 0.11;  
 Matches 12; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Db 13 LLSDSYKREIENGMLWEVGKMWTK 38

RESULT 7  
 T06567  
 IGE-dependent histamine-releasing factor homolog - garden pea  
 N/Alternate names: callus protein p23  
 C/Species: Pisum sativum (garden pea)  
 C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C/Accession: T06567  
 R/Mo, H.H.  
 Submitted to the EMBL Data Library, November 1996  
 A/Description: Primary structure of mRNA encoding a putative 23-kDa callus protein.  
 A/Reference number: Z15764  
 A/Accession: T06567  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-167 <MOO>  
 A/Cross-references: UNIPROT:P50906, EMBL:L47968, NID:91675195, PIND:AAB19090.1, PID:9167





## RESULT 12

E70711

Hypothetical protein Rv1490 - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004

C/Accession: E70711

R/Col: S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentile, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; PMID:98295387; PMID:9634230

A/Accession: E70711

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-435 &lt;COL&gt;

A/Cross-references: UNIPROT:P11771; GB:I279701; GB:AI123456; NID:G3261635; PIDN:CAB02040.

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: Rv1490

C/Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1490

## Query Match

Best Local Similarity 34.7%; Score 50; DB 2; Length 435;

Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 5 IYGIIRIADGLCEVEGKMSRP 27

DB 244 VVGRDIDQLCAMHPHGMENP 266

## RESULT 13

F84054

5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase pfs [imported] - Bacillus halodurans

C/Species: Bacillus halodurans

C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004

C/Accession: F84054

R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A/Reference number: A83650; PMID:10512582; PMID:11058132

A/Accession: F84054

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-233 &lt;STO&gt;

A/Cross-references: UNIPROT:Q9K7X0; GB:AP001518; GB:BA000004; NID:G10175792; PIDN:BA069

A/Experimental source: strain C-125

C/Genetics:

A/Gene: pfs

C/Superfamily: Escherichia coli pfs protein

## Query Match

Best Local Similarity 34.4%; Score 49.5; DB 2; Length 233;

Matches 10; Conservative 9; Mismatches 5; Indels 3; Gaps 1;

QY 3 SDIYGIIRIADGLCEVEGKMSR 26

DB 157 ADHRHRYELAEKPDVCEMGAAYAQ 183

## RESULT 14

A64484

Conserved hypothetical protein MJ1474 [imported] - Methanococcus jannaschii

C/Species: Methanococcus jannaschii

C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004

C/Accession: A64484

R/Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.; Reich, C.I.; Overbeek, R.; Kirsch, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

rsen, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996

A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A/Reference number: A64300; PMID:96337999; PMID:8689087

A/Accession: A64484

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-197 &lt;BUL&gt;

A/Cross-references: UNIPROT:Q58869; GB:U67588; GB:LI77117; NID:G2826411; PIDN:AA099480.1;

C/Genetics:

A/Map position: FOR1445719-1446312

A/Start codon: TTG

C/Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1474

## Query Match

Best Local Similarity 34.0%; Score 49; DB 2; Length 197;

Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 MESDIYGIIRIADGLCEVEG 21

DB 122 LYTDYDGIQVAKKNIIEVKG 142

## RESULT 15

B72725

Hypothetical protein APE0339 - Aeropyrum pernix (strain K1)

C/Species: Aeropyrum pernix

C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C/Accession: B72725

R/Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A/Reference number: A72450; PMID:99310339; PMID:10382966

A/Accession: B72725

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-366 &lt;XAM&gt;

A/Cross-references: UNIPROT:Q9YFA2; DDBJ:AP000059; NID:G5103911; PIDN:BA079294.1; PID:G5

A/Experimental source: strain K1

C/Genetics:

A/Gene: APE0339

C/Superfamily: Aquifex aeolicus hypothetical protein aq\_1142

## Query Match

Best Local Similarity 40.9%; Score 49; DB 2; Length 366;

Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 7 GIREIADGLCEVEGKMSRPE 28

DB 95 GLDIDSSGRCSRCERKHTARPE 116

Search completed: November 10, 2004, 12:29:12  
Job time : 7.66049 secs



RT "cDNA sequence coding for a transcriptionally controlled human tumor  
 RT protein.";  
 RL Nucleic Acids Res. 17:8367-8367(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Thiele H.;  
 RL Thesis (2000), Humboldt-University Berlin, Germany.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung, and Placenta;  
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strassberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buerow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halcy S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 1-10.  
 RC TISSUE=Liver;  
 RX MEDLINE=93162045; PubMed=1286669;  
 RA Hochrasser D.F., Frutiger S., Paquet N., Batroch A., Ravier F.,  
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargias K.,  
 RA Appel R.D., Hughes G.J.;  
 RT "Human liver protein map: a reference database established by  
 RT microsequencing and gel comparison.";  
 RL Electrophoresis 13:992-1001(1992).  
 RN [5]  
 RP SEQUENCE OF 1-4; 22-31; 39-45 AND 103-109.  
 RC TISSUE=Keratinocytes;  
 RX MEDLINE=93162043; PubMed=1286667;  
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,  
 RA Vandekerckhove J.;  
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
 RT protein database of normal human epidermal keratinocytes.";  
 RL Electrophoresis 13:960-969(1992).  
 RN [6]  
 RP SEQUENCE OF 1-19.  
 RC TISSUE=Breast carcinoma;  
 RX MEDLINE=97295304; PubMed=9150946;  
 RA Rasmussen R.K., Ji H., Eddes J.S., Moritz R.L., Reid G.E.,  
 RA Simpson R.J., Dorow D.S.;  
 RT "Two-dimensional electrophoretic analysis of human breast carcinoma  
 RT proteins: mapping of proteins that bind to the SH3 domain of mixed  
 RT lineage kinase MLK2.";  
 RL Electrophoresis 18:588-598(1997).  
 RN [7]  
 RP SEQUENCE OF 1-18.  
 RC MEDLINE=9350660; PubMed=7542803;  
 RA McDonald S.M., Rafnar T., Langdon J., Lichtenstein L.M.;  
 RT "Molecular identification of an Igb-dependent histamine-releasing  
 RT factor.";  
 RL Science 269:688-690(1995).  
 RN [8]  
 RP TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=97213057; PubMed=9059837;  
 RA Sanchez J.-C., Schaller D., Ravier F., Golaz O., Jacoud S., Belet M.,  
 RA Wilkins M.R., James R., Desnusses J., Hochrasser D.F.;  
 RT "Transcriptionally controlled tumor protein: a protein identified in

RT several nonmucosal cells including erythrocytes.";  
 RL Electrophoresis 18:150-155(1997).  
 RN [9]  
 RP PHOSPHORYLATION SITES SER-46 AND SER-64.  
 RX MEDLINE=22157906; PubMed=12167714;  
 RA Yarn F.R.;  
 RT "Plk phosphorylation regulates the microtubule-stabilizing protein  
 RT TCTP.";  
 RL Mol. Cell. Biol. 22:6209-6221(2002).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: Found in several healthy and tumoral cells  
 CC including erythrocytes, hepatocytes, macrophages, platelets,  
 CC keratinocytes, erythroleukemia cells, gliomas, melanomas,  
 CC hepatoblastomas, and lymphomas. It cannot be detected in kidney  
 CC and renal cell carcinoma (RCC).  
 CC -1- SIMILARITY: Belongs to the TCTP family.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X16064; CAA34200.1; -;  
 DR EMBL; AJ400717; CAB87812.1; -;  
 DR EMBL; BC003352; AAH03352.1; -;  
 DR EMBL; BC012431; AAH12431.1; -;  
 DR EMBL; BC052333; AAH52333.1; -;  
 DR FTR; S06590; S06590.  
 DR HSSP; Q10344; 1H6Q.  
 DR SWISS-2DPAGE; P13693; HUMAN.  
 DR Aarhus/Chent-2DPAGE; 8114; IEF.  
 DR Aarhus/Chent-2DPAGE; 9119; IEF.  
 DR OGP; P13693; -;  
 DR Stena-2DPAGE; P13693; -;  
 DR GENE; HGNC:12022; TPT1.  
 DR MIM; 600763; -;  
 DR GO; GO:0005615; Cytoplasmic space; TAS.  
 DR InterPro; IPR011057; Msa4 like.  
 DR InterPro; IPR001983; TCTP.  
 DR Pfam; PF00838; TCTP; 1.  
 DR PRINTS; PRO1653; TCTPROTEIN.  
 DR PRODOM; PD004329; TCTP; 1.  
 DR PROSITE; PS01002; TCTP\_1; 1.  
 DR PROSITE; PS01003; TCTP\_2; 1.  
 KW Direct protein sequencing; Phosphorylation.  
 FT MOD RES 46 46 Phosphoserine (by PLK).  
 FT MOD RES 64 64 Phosphoserine (by PLK).  
 FT CONFLICT 53 54 SA -> YG (in Ref. 3; AAH12431).  
 SQ SEQUENCE 172 AA; 19595 MW; BD31399B71CA62F9 CRC64;  
 QY Query Match 88.9%; Score 128; DB 1; Length 172;  
 Db Best local similarity 92.9%; Pred. No. 6; 1e-11;  
 Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 MFSDYIGREIADGLCEYGVKVSPE 28  
 13 MFSDYIKRIADGLCEYGVKVSPE 40  
 RESULT 3  
 TCTP\_PIG STANDARD; PRT; 172 AA.  
 AC P61286;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Translationally controlled tumor protein (TCTP).  
 GN Name=PRT;  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OK NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yubero N., Barbancho M.J., Llanes D., Garrido J.J.;  
 RT "Molecular characterization of the gene coding for the pig  
 RL translationally controlled tumor protein (TCTP).";  
 CC Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the TCTP family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AY072784; AAL68965.1; -  
 CC PROSITE: PS01002; TCTP\_1; 1.  
 CC PROSITE: PS01003; TCTP\_2; 1.  
 CC Phosphorylation.  
 KW MOD\_RES 46 46 Phosphoserine (by PLX) (By similarity).  
 FT MOD\_RES 64 64 Phosphoserine (by PLX) (By similarity).  
 FT MOD\_RES 64 64 Phosphoserine (by PLX) (By similarity).  
 SQ SEQUENCE 172 AA; 19595 MW; BD31399B71CA62F9 CRC64;

Query Match 88.9%; Score 128; DB 1; Length 172;  
 Best Local Similarity 92.9%; Pred. No. 6,1e-11;  
 Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 MFSDIYGIKRIADGCLCEVGGKMSRPE 28  
 |||||  
 Db 13 MFSDIYKIRIADGCLCEVGGKMSRTE 40

RESULT 4  
 ID Q724U4 PRELIMINARY; PRT; 172 AA.  
 AC Q724U4;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE TCTP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OK NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gnaneskar M., Ramawamy K.; EMBL/GenBank/DBJ databases.  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY34563; AAQ01550.1; -  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR InterPro; IPR01983; TCTP.  
 DR Pfam; PF00838; TCTP; 1.  
 DR ProDom; PD004329; TCTP; 1.  
 DR PROSITE; PS01002; TCTP\_1; 1.  
 DR PROSITE; PS01003; TCTP\_2; 1.  
 SQ SEQUENCE 172 AA; 19594 MW; BD31399B79EA62F9 CRC64;

Query Match 88.9%; Score 128; DB 2; Length 172;  
 Best Local Similarity 92.9%; Pred. No. 6,1e-11;  
 Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 MFSDIYGIKRIADGCLCEVGGKMSRPE 28  
 |||||  
 Db 13 MFSDIYKIRIADGCLCEVGGKMSRTE 40

RESULT 5  
 AAM51565 PRELIMINARY; PRT; 172 AA.  
 ID AAM51565

AC AAM51565;  
 DT 02-MAR-2004 (TREMBLrel. 27, Created)  
 DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)  
 DE P02 protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OK NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gao T.H., Duan F.L., Zhu W.L.;  
 RT "Hepatocarcinoma related gene p02";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY117678; AAM51565.1; -  
 SQ SEQUENCE 172 AA; 19595 MW; BD31399B71CA62F9 CRC64;

Query Match 88.9%; Score 128; DB 2; Length 172;  
 Best Local Similarity 92.9%; Pred. No. 6,1e-11;  
 Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 MFSDIYGIKRIADGCLCEVGGKMSRPE 28  
 |||||  
 Db 13 MFSDIYKIRIADGCLCEVGGKMSRTE 40

RESULT 6  
 ID CAG33317 PRELIMINARY; PRT; 172 AA.  
 AC CAG33317;  
 DT 01-JUN-2004 (TREMBLrel. 27, Created)  
 DT 01-JUN-2004 (TREMBLrel. 27, Last sequence update)  
 DT 01-JUN-2004 (TREMBLrel. 27, Last annotation update)  
 DE TPPI protein.  
 GN TPPI.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OK NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Expert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;  
 RT "Cloning of human full open reading frames in Gateway(TM) system entry  
 RT vector (pDONR201).";  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR457036; CAG33317.1; -  
 SQ SEQUENCE 172 AA; 19595 MW; BD31399B71CA62F9 CRC64;

Query Match 88.9%; Score 128; DB 2; Length 172;  
 Best Local Similarity 92.9%; Pred. No. 6,1e-11;  
 Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 MFSDIYGIKRIADGCLCEVGGKMSRPE 28  
 |||||  
 Db 13 MFSDIYKIRIADGCLCEVGGKMSRTE 40

RESULT 7  
 ID Q862L1 PRELIMINARY; PRT; 114 AA.  
 AC Q862L1;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Similar to tumor protein, translationally-controlled 1  
 DE (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae;  
 OC Bovidae; Bos.  
 OK NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RX MEDLINE=22544902; PubMed=12658628;
RA Ishiwa H., Katsuma S., Kizaki K., Patel O.V., Nakano H.,
RA Takahashi T., Imai K., Hirasawa A., Shiojima S., Ikawa H., Suzuki Y.,
RA Tsujimoto G., Izaike Y., Todoroki J., Hashizume K.;
RT "Characterization of gene expression profiles in early bovine
RT pregnancy using a custom cDNA microarray.";
RL Mol. Reprod. Dev. 65:9-18(2003).
DR EMBL; AB098973; BAC56463.1; -.
DR HSSP; Q10344; 1H6Q.
DR GO; GO:0005737; Cytoplasm; IEA.
DR InterPro; IPR011057; Mes4_1like.
DR InterPro; IPR001983; TCTP-
DR PRINTS; PR01653; TCTPROTEIN.
DR ProDom; PD004329; TCTP; 1.
DR PROSITE; PS01002; TCTP_1; 1.
FT NON_TER 1 114
FT 114
SQ SEQUENCE 114 AA; 12879 MW; 2D2A5CE73183FEDE CRC64;

Query Match 88.2%; Score 127; DB 2; Length 114;
Best Local Similarity 89.3%; Pred. No. 5.8e-11;
Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MFSDIYKIREVADGTCLEVEGKMSRPE 28
Db 13 MFSDIYKIREVADGTCLEVEGKMSRTE 40

RESULT 8
Q862H4 PRELIMINARY; PRT; 117 AA.
ID Q862H4;
AC Q862H4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to tumor protein, translationally-controlled 1
DE (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22544902; PubMed=12658628;
RA Ishiwa H., Katsuma S., Kizaki K., Patel O.V., Nakano H.,
RA Takahashi T., Imai K., Hirasawa A., Shiojima S., Ikawa H., Suzuki Y.,
RA Tsujimoto G., Izaike Y., Todoroki J., Hashizume K.;
RT "Characterization of gene expression profiles in early bovine
RT pregnancy using a custom cDNA microarray.";
RL Mol. Reprod. Dev. 65:9-18(2003).
DR EMBL; AB099016; BAC56506.1; -.
DR HSSP; Q10344; 1H6Q.
DR GO; GO:0005737; Cytoplasm; IEA.
DR InterPro; IPR011057; Mes4_1like.
DR InterPro; IPR001983; TCTP-
DR PRINTS; PR01653; TCTPROTEIN.
DR ProDom; PD004329; TCTP; 1.
DR PROSITE; PS01002; TCTP_1; 1.
FT NON_TER 1 117
FT 117
SQ SEQUENCE 117 AA; 13192 MW; B22021FDC8E7731 CRC64;

Query Match 88.2%; Score 127; DB 2; Length 117;
Best Local Similarity 89.3%; Pred. No. 6e-11;
Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MFSDIYKIREVADGTCLEVEGKMSRPE 28
Db 13 MFSDIYKIREVADGTCLEVEGKMSRTE 40

```

```

Q862G3 PRELIMINARY; PRT; 155 AA.
ID Q862G3;
AC Q862G3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to translationally controlled tumor protein (fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22544902; PubMed=12658628;
RA Ishiwa H., Katsuma S., Kizaki K., Patel O.V., Nakano H.,
RA Takahashi T., Imai K., Hirasawa A., Shiojima S., Ikawa H., Suzuki Y.,
RA Tsujimoto G., Izaike Y., Todoroki J., Hashizume K.;
RT "Characterization of gene expression profiles in early bovine
RT pregnancy using a custom cDNA microarray.";
RL Mol. Reprod. Dev. 65:9-18(2003).
DR EMBL; AB099031; BAC56521.1; -.
DR HSSP; Q10344; 1H6Q.
DR GO; GO:0005737; Cytoplasm; IEA.
DR InterPro; IPR011057; Mes4_1like.
DR InterPro; IPR001983; TCTP-
DR PRINTS; PR01653; TCTPROTEIN.
DR ProDom; PD004329; TCTP; 1.
DR PROSITE; PS01003; TCTP_2; 1.
FT NON_TER 1 155
FT 155
SQ SEQUENCE 155 AA; 17548 MW; B8EBD7E5DAF21D6 CRC64;

Query Match 88.2%; Score 127; DB 2; Length 155;
Best Local Similarity 89.3%; Pred. No. 7.8e-11;
Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MFSDIYKIREVADGTCLEVEGKMSRPE 28
Db 13 MFSDIYKIREVADGTCLEVEGKMSRTE 40

RESULT 10
TCTP_MOUSE STANDARD; PRT; 172 AA.
ID TCTP_MOUSE;
AC P14701;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Translationally controlled tumor protein (TCTP) (p23) (21 kDa
DE polypeptide) (p21) (Lens epithelial protein).
GN Name=tpst; Synonyms=Itc;
OS Mus musculus (Mouse); and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090, 10116;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=Mouse;
RX MEDLINE=88189840; PubMed=3357792;
RX Makrides S., Chitpathira S.T., Bandopadhyay R., Braerman G.;
RT "Nucleotide sequence of a major messenger RNA for a 21 kilodalton
RT polypeptide that is under translational control in mouse tumor
RT cells.";
RL Nucleic Acids Res. 16:2350-2350(1988).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX SPECIES=Mouse;
RX MEDLINE=90056510; PubMed=2479380;
RX Boehm H., Beendorf R., Gaestel M., Gross B., Nuernberg P., Kraft R.,
RX Otto A., Bielek H.;

```

```

RT "The growth-related protein P23 of the Ehrlich ascites tumor:
RT translational control, cloning and primary structure.",
RT Biochem. Int. 19:277-286(1989).
RN [3]
RN SEQUENCE FROM N.A.
RC SPECIES=rat; STRAIN=Sprague-Dawley; TISSUE=lens;
RA Wen Y., Li G., Chen P., Bekhor I.,
RA Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RN INTERACTIONS.
RP MEDLINE=21260623; PubMed=11368327;
RX Yoon T., Jung J., Kim M., Lee K.M., Choi E.C., Lee K.;
RT "Identification of the self-interaction of rat TCTP/19E-dependent
RT histamine-releasing factor using yeast two-hybrid system.";
RL Arch. Biochem. Biophys. 384:379-382(2000).
RN [5]
RN CALCULUM-BINDING.
RP MEDLINE=21027882; PubMed=1156187;
RX Kim M., Jung Y., Lee K., Kim C.;
RT "Identification of the calcium binding sites in translationally
RT controlled tumor protein.";
RL Arch. Pharm. Res. 23:633-636(2000).
CC -1- SUBUNIT: Seems to self-interact.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- DEVELOPMENTAL STAGE: Preferentially synthesized in cells of the
CC early growth phase of Ehrlich ascites tumor.
CC -1- SIMILARITY: Belongs to the TCTP family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06407; CAA29697.1; -
DR EMBL; U20525; AAA62507.1; -
DR PIR; S00775; S00775.
DR HSRF; Q10344; HRGQ.
DR SWISS-2DPAGE; P14701; MOUNS.
DR PMMA-2DPAGE; P14701; -.
DR MGD; MGI:104890; Tpt1.
DR RGD; 621623; Tpt1.
DR InterPro; IPR011057; Msa4_like.
DR InterPro; IPR001983; TCTP.
DR Pfam; PF00638; TCTP; 1.
DR PRINTS; PR01653; TCTPROTEIN.
DR PRODOM; PD004329; TCTP; 1.
DR PROSITE; PS01002; TCTP; 1.
DR PROSITE; PS01003; TCTP; 2; 1.
KW Calcium-binding; Direct protein sequencing; Phosphorylation.
FT MOD_RES 46 46 Phosphoserine (by PLK) (By similarity).
FT MOD_RES 64 64 Phosphoserine (by PLK) (By similarity).
SQ SEQUENCE 172 AA; 19462 MW; 7F46751DA2D281D2 CRC64;

```

```

GN Tpt1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=129/Sv;
RX MEDLINE=22317404; PubMed=12399545;
RA Tyndler M., Susini L., Prieur S., Besse S., Flucci G., Amson R.,
RA Teلمان A.;
RT "Biological models and genes of tumor reversion: cellular
RT reprogramming through tpt1/tctp and SIM-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14976-14981(2002).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=129/Sv;
RA Flucci G., Lespagnol A., Stumpfner-Cveleto P., Beaucourt S.,
RA Dulac D., Susini L., Amson R., Teلمان A.;
RT "Genomic organization and expression of mouse tpt1 gene.";
RL Genomics 81:570-578(2003).
DR EMBL; AY186881; AAP23875.1; -
SQ SEQUENCE 172 AA; 19462 MW; 7F46751DA2D281D2 CRC64;

Query Match 86.8%; Score 125; DB 2; Length 172;
Best Local Similarity 89.3%; Pred. No. 1.7e-10;
Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 MFSDYIGRIADGCLCEVEGKMSRPE 28
Db 13 LPSDIYKIRIADGCLCEVEGKMSRTE 40

RESULT 12
TCTP RABIT STANDARD; PRT; 172 AA.
AC PA3348; 077730; Q9MYT9;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Translationally controlled tumor protein (TCTP).
GN Name=TPT1;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Mammary gland;
RA Dawson S.P., Martin M., Tighe P.J., Wilde C.J., Mayer R.J.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=99013447; PubMed=9799103;
RA Thiele H., Berger M., Ienzer C., Kuhn H., Thiele B.-J.;
RT "Structure of the promoter and complete sequence of the gene coding
RT for the rabbit translationally controlled tumor protein (TCTP) P23.";
RL Eur. J. Biochem. 257:62-68(1998).
RN [3]
RN SEQUENCE FROM N.A., AND VARIANTS.
RX MEDLINE=20408860; PubMed=10951206;
RA Thiele H., Berger M., Skaliweit A., Thiele B.-J.;
RT "Expression of the gene and processed pseudogenes encoding the human
RT and rabbit translationally controlled tumor protein (TCTP).";
RL Eur. J. Biochem. 267:5473-5481(2000).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- DEVELOPMENTAL STAGE: Undergoes developmental regulation during
CC mammary gland development.
CC -1- SIMILARITY: Belongs to the TCTP family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; 246805; CAA68260.1; -  
DR EMBL; AJ228598; CAA12650.1; -  
DR EMBL; AJ131951; CAB41990.1; -  
DR EMBL; AJ277093; CAC01238.1; -  
DR EMBL; AJ277093; CAC01240.1; -  
DR PIR; A38956; A38956.  
DR HSP; Q10344; 1H6Q.  
DR InterPro; IPR011057; Mss4\_1like.  
DR InterPro; IPR011983; TCTP-  
DR Pfam; PF00838; TCTP; 1.  
DR PRINTS; PR01653; TCTPROTEIN.  
DR PRODOM; PD004329; TCTP; 1.  
DR PROSITE; PS01002; TCTP\_1; 1.  
DR PROSITE; PS01003; TCTP\_2; 1.  
KW Phosphorylation.  
FT MOD RES 46 46 Phosphoserine (by PLK) (By similarity).  
FT MOD RES 64 64 Phosphoserine (by PLK) (By similarity).  
FT VARIANT 79 79 O -> R.  
FT VARIANT 153 153 E -> G.  
FT VARIANT 169 169 Y -> V.  
FT VARIANT 169 169 Y -> D (in Ref. 1).  
FT CONFLICT 13 13 M -> T (in Ref. 1).  
FT CONFLICT 129 129 F -> L (in Ref. 1).  
SQ SEQUENCE 172 AA; 19537 MW; 7E3209A92C601452 CRC64;

Query Match 84.0%; Score 121; DB 1; Length 172;  
Best Local Similarity 89.3%; Pred. No. 7e-10; 3; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MFSDYIGIREADGCLVEYEGMVSPE 28  
Db 13 MFSDYIKIREANGCLVEYEGMVSPE 40

RESULT 13  
TCTP\_CHICK STANDARD; PRT; 172 AA.  
ID TCTP\_CHICK  
AC P43347;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 05-UTL-2004 (Rel. 44, Last annotation update)  
DE Translationally controlled tumor protein (TCTP) (p23) (pCHK23).  
GN Name=TPT1;  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
NCBI\_TaxID=9031;  
RN NCBI  
RP SEQUENCE FROM N.A.  
RC TISSUE=Myoblasts;  
RA Kang B.-S., Yang S., Yoo B., Yang J.;  
RT Cloning and nucleotide sequence of differentiation-related gene  
RT (pCHK23) from chick embryonic myoblasts; cDNA pCHK23 is homologous to  
RT translationally controlled protein p23 in mouse ERI cells.  
RL Submitted (APR-1995) to the EMBL/GenBank/DBS databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lens;  
RA Sawada K., Agata K., Eguchi G.;  
RL Submitted (DEC-1993) to the EMBL/GenBank/DBS databases.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the TCTP family.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U25954; AAA67296.1; -  
DR EMBL; D26312; BAA05374.1; -  
DR PIR; A38960; A38960.  
DR HSP; Q10344; 1H6Q.  
DR InterPro; IPR011057; Mss4\_1like.  
DR InterPro; IPR011983; TCTP-  
DR Pfam; PF00838; TCTP; 1.  
DR PRINTS; PR01653; TCTPROTEIN.  
DR PRODOM; PD004329; TCTP; 1.  
DR PROSITE; PS01002; TCTP\_1; 1.  
DR PROSITE; PS01003; TCTP\_2; 1.  
KW Phosphorylation.  
FT MOD RES 46 46 Phosphoserine (by PLK) (By similarity).  
FT MOD RES 129 129 F -> L (in Ref. 1).  
SQ SEQUENCE 172 AA; 19530 MW; 3D3DB669AC8E58F1 CRC64;

Query Match 82.6%; Score 119; DB 1; Length 172;  
Best Local Similarity 82.1%; Pred. No. 1.4e-09; 2; Indels 0; Gaps 0;  
Matches 23; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MFSDYIGIREADGCLVEYEGMVSPE 28  
Db 13 MFSDYIKIREANGCLVEYEGMVSPE 40

RESULT 14  
Q7ZYF2 PRELIMINARY; PRT; 172 AA.  
ID Q7ZYF2;  
AC Q7ZYF2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Tpt1-prov protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodidae; Xenopus.  
NCBI\_TaxID=8355;  
RN NCBI  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Vax S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Muliyil S.J.,  
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,  
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman W., Madan A.C., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzyzanski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,  
RA Jones S.V., Maizra M.A.;  
RT "generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917;  
RA Klein S.U., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Rasmussen P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus



```

RT      initiaive."
RL      Dev. Dyn. 225:384-391(2002).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Embryo;
RA      Klein S., Strausberg R.;
RL      Submitted (Jan-2003) to the EMBL/GenBank/DBJ databases
DR      EMBL; BC043811; AAH43811.1; -.
DR      HSSP; Q10344; 1H69.
DR      GO; GO:005573; C:cycloplasm; IPA.
DR      InterPro; IPR011057; Mss4_like.
DR      InterPro; IPR011983; TCIP.
DR      PRINTS; PR01653; TCIPROTEIN.
DR      ProDom; PD004329; TCIP; 1.
DR      PROSITE; PS01003; TCIP; 2; 1.
SQ      SEQUENCE 172 AA, 19526 MW, 1F34F2B7B1E9D168 CRC64;

```

Query Match	68.8%	Score 99	DB 2	Length 172
Best Local Similarity	67.9%	Pred. No. 1.4e-06		
Matches 19	Conservative 3	Mismatches 6	Indels 0	Gaps 0

```

QY      1 MFSDIYGIREFIADGLCLEVEGKMSRPE 28
          |||||  |||:|||||:|
Db      13 MFSDIYKIIFTPDGMCLVEGKVIQREE 40

```

RESULT 15

AC	ORGANISM	PRELIMINARY	PRT	140 AA.
AD	O62T11			
AD	O62T11			
DT	05-JUL-2004 (TrEMBLrel. 27, Created)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)			
DE	Hypothetical protein FLJ44635.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Cerebellum;			
RA	Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hirao M., Shimizu F.,			
RA	Wakabe H., Ono T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T.,			
RA	Irie R., Otsuki T., Sato H., Wakamatsu A., Ishi S., Yamamoto J.,			
RA	Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,			
RA	Yanashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,			
RA	Wagatsuma M., Morikawa K., Kanehori K., Takahashi-Fuji A., Oshima A.,			
RA	Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nabehari K.,			
RL	Submitted (JUL-2003) to the EMBL/GenBank/DSJ databases.			
DR	EMBL: AK126598; BAC6606.1; -;			
DR	InterPro: IPR011057; Meta4_like.			
DR	InterPro: IPR001983; TCTP.			
DR	Pfam: PF00838; TCTP. 1.			
DR	PRINTS: PRO1653; TCTPROTEIN.			
DR	ProDom: PD004329; TCTP. 1.			
DR	PROSITE: PS01002; TCTP. 1, 1.			
DR	SEQUENCE 140 AA; 15757 MW; 1ADEB7A0C9866663 CRC64;			
SC	SEQUENCE 140 AA; 15757 MW; 1ADEB7A0C9866663 CRC64;			

Query Match	66.0%;	Score 95;	DB 2;	Length 140;
Best Local Similarity	75.0%;	Pred. No. 4.8e-06;		
Matches	21;	Mismatches	5;	Indels 0;
		Conservative		Gaps 0;

```

Oy      1 MFSDIYGIREFIADGLCLVEEGKMWSPRE 28
          |||:|||||:|||||
Db      18 MFSDSYMSQEIADGLRLVEEGKIVSRTE 45

```

```
Search completed: November 10, 2004, 12:27:01
Job time : 31.6358 secs
```

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 ; Search time 7.44599 Seconds  
(without alignments)  
222.664 Million cell updates/sec

Title: US-10-092-750-11

Perfect score: 127

Sequence: 1 FWLESRDFEAGVFLEAIVNSIKRS 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 segs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	46.5	377	4	US-09-107-532A-4318
2	47.5	37.4	676	4	US-09-583-110-4259
3	47	37.0	94	4	US-09-248-796A-18013
4	47	37.0	482	2	US-08-724-194-4
5	47	37.0	482	2	US-08-724-194-5
6	47	37.0	482	2	US-08-555-758C-2
7	46	36.2	538	2	US-08-209-521-6
8	46	36.2	538	3	US-08-961-810-2
9	46	36.2	538	3	US-08-352-902D-2
10	46	36.2	538	4	US-09-265-503B-2
11	46	36.2	649	3	US-08-676-444-43
12	46	36.2	649	4	US-09-583-110-4809
13	45.5	35.8	440	4	US-09-489-039A-10762
14	45	35.4	194	4	US-09-543-681A-8016
15	45	35.4	648	4	US-09-328-352-4885
16	45	35.4	784	4	US-09-982-308B-23
17	45	35.4	968	1	US-08-434-730-14
18	45	35.4	1185	3	US-08-664-962B-2
19	45	35.4	1185	3	US-09-311-743-2
20	44.5	35.0	336	4	US-09-248-796A-17023
21	44	34.6	431	4	US-09-107-532A-7056
22	44	34.6	431	4	US-09-134-000C-6516
23	44	34.6	465	4	US-09-134-000C-6425
24	44	34.6	481	3	US-09-398-165-2
25	44	34.6	481	3	US-08-948-559-2
26	43.5	34.3	157	4	US-09-071-035-328
27	43.5	34.3	264	4	US-10-101-464A-959

28	43.5	34.3	851	4	US-09-071-035-326	Sequence 326, App
29	43.5	34.3	851	4	US-09-071-035-330	Sequence 330, App
30	43.5	34.3	851	4	US-09-071-035-334	Sequence 334, App
31	43.5	34.3	962	4	US-09-134-000C-4457	Sequence 4257, App
32	43	33.9	62	4	US-09-248-796A-25821	Sequence 25821, A
33	43	33.9	64	4	US-09-107-532A-4577	Sequence 4577, App
34	43	33.9	80	4	US-09-328-352-5195	Sequence 5195, App
35	43	33.9	284	4	US-09-252-991A-26771	Sequence 26771, A
36	43	33.9	378	4	US-09-134-000C-4804	Sequence 4804, App
37	42	33.1	70	4	US-09-583-110-4603	Sequence 4603, App
38	42	33.1	112	4	US-09-543-681A-5924	Sequence 5924, App
39	42	33.1	149	4	US-09-445-160B-9	Sequence 9, Appli
40	42	33.1	156	4	US-09-270-767-35281	Sequence 35281, A
41	42	33.1	156	4	US-09-270-767-50498	Sequence 50498, A
42	42	33.1	189	4	US-09-248-796A-17133	Sequence 17133, A
43	42	33.1	315	3	US-09-134-001C-3786	Sequence 3786, App
44	42	33.1	356	4	US-09-107-532A-3683	Sequence 3683, App
45	42	33.1	391	4	US-09-489-039A-7804	Sequence 7804, App

## ALIGNMENTS

RESULT 1  
US-09-107-532A-4318  
Sequence 4318, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID and AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSER: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
Application Number: US/09/107,532A  
Filing Date: 30-Jun-1998  
Prior Application Data:  
Application Number: 60/085,598  
Filing Date: 14 May 1998  
Application Number: 60/051571  
Filing Date: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Artniello, Pamela Danek  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 4318:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURES:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...377  
SEQUENCE DESCRIPTION: SEQ ID NO: 4318:  
US-09-107-532A-4318

Query Match 46.5%; Score 59; DB 4; Length 377;  
Best Local Similarity 37.5%; Pred. No. 0.17;  
Matches 9; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 2 WLEERDFEAGVFELEAIVNSIKR 25  
Db 133 WKERAYEAGIRPVDLITSAXKS 156

RESULT 2  
US-09-583-110-4259  
Sequence 4259, Application US/09583110  
Patent No. 6699703  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al.  
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
FILE REFERENCE: P4700-07A  
CURRENT APPLICATION NUMBER: US/09/583,110  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/107,433  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/085,131  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: US 60/051,553  
PRIOR FILING DATE: 1997-07-02  
NUMBER OF SEQ ID NOS: 5332  
SEQ ID NO 4259  
LENGTH: 676  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-583-110-4259

Query Match 37.4%; Score 47.5; DB 4; Length 676;  
Best Local Similarity 50.0%; Pred. No. 26;  
Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

Qy 6 RDF---EAGVFELEAIVNSIKR 24  
Db 478 KDFPNLEYGDFEIEQILINIVKR 499

RESULT 3  
US-09-248-796A-18013  
Sequence 18013, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 18013  
LENGTH: 94  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-18013

Query Match 37.0%; Score 47; DB 4; Length 94;  
Best Local Similarity 45.5%; Pred. No. 3.1;  
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 3 LEERDFEAGVFELEAIVNSIKR 24  
Db 31 LEEVSFNLCGLDLDQIVSVGR 52

RESULT 4  
US-08-724-194-4  
Sequence 4, Application US/08724194  
Patent No. 5824875  
GENERAL INFORMATION:  
APPLICANT: RANU, RAJINDER S.  
TITLE OF INVENTION: ONE-AMINOCYCLOPROPANE-1-CARBOXYLATE  
TITLE OF INVENTION: SYNTHASE GENES FROM PELARGONIUM TO CONTROL ETHYLENE LEVELS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSER: SANTANGELO LAW OFFICES PC  
STREET: 315 WEST OAK STREET, STE 701  
CITY: FORT COLLINS  
STATE: CO  
COUNTRY: USA  
ZIP: 80521  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,194  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: SANTANGELO, LUKE  
REGISTRATION NUMBER: 31,997  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (970) 224-3100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 482 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-724-194-4

Query Match 37.0%; Score 47; DB 2; Length 482;  
Best Local Similarity 37.9%; Pred. No. 21;  
Matches 11; Conservative 4; Mismatches 8; Indels 6; Gaps 1;

Qy 1 FW-----LEERDFEAGVFELEAIVNSIK 23  
Db 363 FWDLRKLEBEETFEAMVLMKYIINEVK 391

RESULT 5  
US-08-724-194-5  
Sequence 5, Application US/08724194  
Patent No. 5824875  
GENERAL INFORMATION:  
APPLICANT: RANU, RAJINDER S.  
TITLE OF INVENTION: ONE-AMINOCYCLOPROPANE-1-CARBOXYLATE  
TITLE OF INVENTION: SYNTHASE GENES FROM PELARGONIUM TO CONTROL ETHYLENE LEVELS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSER: SANTANGELO LAW OFFICES PC  
STREET: 315 WEST OAK STREET, STE 701  
CITY: FORT COLLINS  
STATE: CO  
COUNTRY: USA  
ZIP: 80521  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,194

FILING DATE: 01-OCT-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: SANTIANGELO, LUKE  
REGISTRATION NUMBER: 31,997  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (970) 224-3100  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 482 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-724-194-5

Query Match 37.0%; Score 47; DB 2; Length 482;  
Best Local Similarity 37.9%; Pred. No. 21;  
Matches 11; Conservative 4; Mismatches 8; Indels 6; Gaps 1;

QY 1 FW-----LEERDFEAGVFELEATVNSIK 23  
DB 363 FWMIDRKLEETFEAEVMTKVIINEVK 391

RESULT 6  
US-08-555-755C-2  
Sequence 2, Application US/08555755C  
Patent No. 6610909  
GENERAL INFORMATION:  
APPLICANT: Oglevee-O'Donovan, Wendy  
APPLICANT: Arteca, Richard N  
APPLICANT: Arteca, Jeanette  
APPLICANT: Stoots, Eleanor  
TITLE OF INVENTION: METHOD FOR THE COMMERCIAL PRODUCTION OF TRANSGENIC  
TITLE OF INVENTION: PLANTS  
FILE REFERENCE: 176-930539  
CURRENT APPLICATION NUMBER: US/08/555,755C  
CURRENT FILING DATE: 1995-11-09  
PRIOR APPLICATION NUMBER: 08/149,702  
PRIOR FILING DATE: 1993-11-09  
PRIOR APPLICATION NUMBER: 07/690,073  
PRIOR FILING DATE: 1991-04-23  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 482  
TYPE: PRT  
ORGANISM: Pelargonium x hortorum  
US-08-555-755C-2

Query Match 37.0%; Score 47; DB 4; Length 482;  
Best Local Similarity 37.9%; Pred. No. 21;  
Matches 11; Conservative 4; Mismatches 8; Indels 6; Gaps 1;

QY 1 FW-----LEERDFEAGVFELEATVNSIK 23  
DB 363 FWMIDRKLEETFEAEVMTKVIINEVK 391

RESULT 7  
US-08-209-521-6  
Sequence 6, Application US/08209521  
Patent No. 5922855  
GENERAL INFORMATION:  
APPLICANT: Liskey, Robert M.  
APPLICANT: Bronner, C. Eric  
APPLICANT: Baker, Sean M.  
APPLICANT: Bollag, Roni J.  
APPLICANT: Kolodner, Richard D.  
TITLE OF INVENTION: MAMMALIAN DNA MISMATCH REPAIR GENES  
TITLE OF INVENTION: HMHI AND HMPS1  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &  
ADDRESSEE: Heuser  
STREET: 520 S.W. Yamhill, Suite 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: US  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,521  
FILING DATE: 08-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysselberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: CHSU 306A  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 538 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-209-521-6

Query Match 36.2%; Score 46; DB 2; Length 538;  
Best Local Similarity 26.2%; Pred. No. 35;  
Matches 11; Conservative 7; Mismatches 6; Indels 18; Gaps 1;

QY 2 WLEERDFEAGVFELEATVNSIKS 25  
DB 458 WMAEBEISGIYEMCDMLLTKEVSIKRYRAELATVNSICKS 499

RESULT 8  
US-08-961-810-2  
Sequence 2, Application US/08961810  
Patent No. 6165713  
GENERAL INFORMATION:  
APPLICANT: Liskey, Robert M.  
APPLICANT: Bronner, C. Eric  
APPLICANT: Baker, Sean M.  
APPLICANT: Bollag, Roni J.  
APPLICANT: Kolodner, Richard D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA  
TITLE OF INVENTION: MISMATCH REPAIR GENES  
NUMBER OF SEQUENCES: 134  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &  
ADDRESSEE: Heuser  
STREET: 520 S.W. Yamhill Street, Suite 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: U.S.A.  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,810  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysselberghe, Pierre C.

REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
TELEX: 360619  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 538 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-961-810-2

Query Match 36.2%; Score 46; DB 3; Length 538;  
Best Local Similarity 26.2%; Pred. No. 35;  
Matches 11; Conservative 7; Mismatches 6; Indels 18; Gaps 1;

QY 2 WLEERDFEAGVPEL-----EATVNSIKRS 25  
DB 458 WMAEEIESGIYEMCDMLLTKEVSIKYRAELAIMSCRS 499

RESULT 9  
US-08-352-902D-2  
Sequence 2, Application US/08352902D  
Patent No. 6191268  
GENERAL INFORMATION:  
APPLICANT: Liskay, Robert M.  
Bromner, C. Eric  
Baker, Sean M.  
Bollag, Roni J.  
Kolodner, Richard D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA  
NUMBER OF SEQUENCES: 149  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack & Heuser  
STREET: 520 S.W. Yamhill Street, Suite 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: U.S.A.  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/352,902D  
FILING DATE: 09-Dec-1994  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysseberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
TELEX: 360619  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 538 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-08-352-902D-2

Query Match 36.2%; Score 46; DB 3; Length 538;

Best Local Similarity 26.2%; Pred. No. 35;  
Matches 11; Conservative 7; Mismatches 6; Indels 18; Gaps 1;  
QY 2 WLEERDFEAGVPEL-----EATVNSIKRS 25  
DB 458 WMAEEIESGIYEMCDMLLTKEVSIKYRAELAIMSCRS 499

RESULT 10  
US-09-265-503B-2  
Sequence 2, Application US/09265503B  
Patent No. 6538108  
GENERAL INFORMATION:  
APPLICANT: Liskay, Robert M.  
Bromner, C. Eric  
Baker, Sean M.  
APPLICANT: Bollag, Roni J.  
APPLICANT: Kolodner, Richard D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS  
RELATING TO DNA MISMATCH REPAIR GENES  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack & Heuser  
STREET: 520 S.W. Yamhill Street, Suite 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: U.S.A.  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/265,503B  
FILING DATE: March 10, 1999  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysseberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
TELEX: 360619  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 538 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-265-503B-2

Query Match 36.2%; Score 46; DB 4; Length 538;  
Best Local Similarity 26.2%; Pred. No. 35;  
Matches 11; Conservative 7; Mismatches 6; Indels 18; Gaps 1;

QY 2 WLEERDFEAGVPEL-----EATVNSIKRS 25  
DB 458 WMAEEIESGIYEMCDMLLTKEVSIKYRAELAIMSCRS 499

RESULT 11  
US-08-676-444-43  
Sequence 43, Application US/08676444A  
Patent No. 6294325  
GENERAL INFORMATION:  
APPLICANT: Wetmur, James G.  
TITLE OF INVENTION: CLONING AND EXPRESSION OF THERMOSTABLE  
NUTL GENES AND PROTEINS AND USES THEREFOR  
FILE REFERENCE: WSM95-02  
CURRENT APPLICATION NUMBER: US/08/676,444A



*This page blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 ; Search time 24.1127 Seconds  
(without alignments)  
366.225 Million cell updates/sec

Title: US-10-092-750-11  
Perfect score: 127  
Sequence: 1 FWLEERDFEAGVFELIEAIVNSIKRS 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225866 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US08A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	100.0	25	US-10-092-750-11	Sequence 11, Appl
2	114	89.8	904	US-09-950-041-6	Sequence 6, Appl
3	114	89.8	904	US-10-265-072-2	Sequence 2, Appl
4	114	89.8	904	US-10-407-952-0	Sequence 2, Appl
5	114	89.8	904	US-10-733-563-6	Sequence 6, Appl
6	114	89.8	904	US-10-733-796A-6	Sequence 6, Appl
7	101	79.5	905	US-10-265-072-4	Sequence 4, Appl
8	101	79.5	905	US-10-407-952-22	Sequence 2, Appl
9	59	46.5	367	US-10-282-122A-57969	Sequence 57969, A
10	50	39.4	117	US-10-424-599-149407	Sequence 149407, A
11	49	38.6	925	US-10-369-493-5415	Sequence 5415, Ap
12	48	37.8	651	US-10-282-122A-72448	Sequence 72448, A
13	48	37.8	713	US-10-406-666A-23	Sequence 23, Appl

14	47.5	37.4	666	9	US-09-815-242-13637	Sequence 13637, A
15	47.5	37.4	676	9	US-09-815-242-13513	Sequence 13513, A
16	47.5	37.4	676	15	US-10-282-122A-74215	Sequence 74215, A
17	47	37.0	144	15	US-10-296-115-1182	Sequence 1182, Ap
18	47	37.0	171	10	US-09-950-041-16	Sequence 18, Appl
19	47	37.0	350	10	US-09-906-179A-21	Sequence 21, Appl
20	47	37.0	1032	10	US-09-950-041-37	Sequence 37, Appl
21	47	37.0	1040	9	US-09-864-761-38325	Sequence 38325, A
22	47	37.0	1041	9	US-09-168-978-3	Sequence 3, Appl1
23	47	37.0	1041	9	US-09-978-697-498	Sequence 498, App
24	47	37.0	1041	9	US-09-978-697-498	Sequence 498, App
25	47	37.0	1041	9	US-09-978-697-498	Sequence 498, App
26	47	37.0	1041	9	US-09-978-697-498	Sequence 498, App
27	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
28	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
29	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
30	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
31	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
32	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
33	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
34	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
35	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
36	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
37	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
38	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
39	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
40	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
41	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
42	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
43	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
44	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App
45	47	37.0	1041	10	US-09-978-697-498	Sequence 498, App

## ALIGNMENTS

RESULT 1  
US-10-092-750-11  
Sequence 11, Application US/10092750  
Publication No. US2003032157A1  
GENERAL INFORMATION:  
APPLICANT: Hammond, Philip W.  
APPLICANT: Alpin, Julia  
TITLE OF INVENTION: Polypeptides Interactive with BCL-X1  
FILE REFERENCE: 50036/050002  
CURRENT FILING DATE: 2002-03-07  
PRIORITY FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: US 60/274,526  
NUMBER OF SEQ ID NOS: 253  
SOFTWARE: PatsSeq for Windows Version 4.0  
SEQ ID NO 11  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-092-750-11

Query Match 100.0%; Score 127; DB 14; Length 25;  
Best Local Similarity 100.0%; Pred. No. 66-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FWLEERDFEAGVFELIEAIVNSIKRS 25  
Db 1 FWLEERDFEAGVFELIEAIVNSIKRS 25

RESULT 2  
US-09-950-041-6  
Sequence 6, Application US/09950041  
Publication No. US2003032090A1  
GENERAL INFORMATION:

APPLICANT: Hardiman, Gerard T.  
APPLICANT: Rock, Fernando L.  
APPLICANT: Bazan, J. Fernando  
APPLICANT: Kastelein, Robert A.  
APPLICANT: Ho, Stephen W.K.  
APPLICANT: Liu, Yong-Jun  
TITLE OF INVENTION: RECEPTOR PROTEINS; RELATED REAGENTS AND METHODS  
FILE REFERENCE: DX0724XKI  
CURRENT APPLICATION NUMBER: US/09/950,041  
CURRENT FILING DATE: 2002-05-06  
PRIOR APPLICATION NUMBER: 09/728,540  
PRIOR FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: 60/207,558  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: 09/073,363  
PRIOR FILING DATE: 1999-06-05  
PRIOR APPLICATION NUMBER: 60/044,293  
PRIOR FILING DATE: 1997-05-07  
PRIOR APPLICATION NUMBER: 60/072,212  
PRIOR FILING DATE: 1998-01-22  
PRIOR APPLICATION NUMBER: 60/076,947  
PRIOR FILING DATE: 1998-03-05  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 6  
LENGTH: 904  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-950-041-6

Query Match 89.8%; Score 114; DB 10; Length 904;  
Best Local Similarity 96.0%; Pred. No. 4,1e-09;  
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FWLEERDFEAGVFELEAIYNSIKRS 25  
DB 786 FCLERDFEAGVFELEAIYNSIKRS 810

RESULT 3  
US-10-265-072-2  
Sequence 2, Application US/10265072  
Publication No. US20030166001A1  
GENERAL INFORMATION:  
APPLICANT: Lipford, Grayson  
TITLE OF INVENTION: TOLL-LIKE RECEPTOR 3 SIGNALING AGONISTS AND ANTAGONISTS  
FILE REFERENCE: CO1041,70031,US  
CURRENT APPLICATION NUMBER: US/10/265,072  
CURRENT FILING DATE: 2002-10-05  
NUMBER OF SEQ ID NOS: 117  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 904  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-265-072-2

Query Match 89.8%; Score 114; DB 14; Length 904;  
Best Local Similarity 96.0%; Pred. No. 4,1e-09;  
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FWLEERDFEAGVFELEAIYNSIKRS 25  
DB 786 FCLERDFEAGVFELEAIYNSIKRS 810

US-10-407-952-20  
Sequence 20, Application US/10407952  
Publication No. US20030232074A1  
GENERAL INFORMATION:  
APPLICANT: Lipford, Grayson  
APPLICANT: Bauer, Stefan

TITLE OF INVENTION: Immunostimulatory G,U-Containing Oligoribonucleotides  
FILE REFERENCE: CO1041,70037,US  
CURRENT APPLICATION NUMBER: US/10/407,952  
CURRENT FILING DATE: 2003-04-04  
PRIOR APPLICATION NUMBER: US 60/421,966  
PRIOR FILING DATE: 2002-10-29  
PRIOR APPLICATION NUMBER: US 60/370,515  
PRIOR FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 20  
LENGTH: 904  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-407-952-20

Query Match 89.8%; Score 114; DB 14; Length 904;  
Best Local Similarity 96.0%; Pred. No. 4,1e-09;  
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FWLEERDFEAGVFELEAIYNSIKRS 25  
DB 786 FCLERDFEAGVFELEAIYNSIKRS 810

RESULT 5  
US-10-732-563-6  
Sequence 6, Application US/10732563  
Publication No. US20040132079A1  
GENERAL INFORMATION:  
APPLICANT: Gupta, Shalley K.  
APPLICANT: Ghosh, Tarun K.  
APPLICANT: Pink, Jason R.  
TITLE OF INVENTION: Assays Relating to Toll-Like Receptor Activity  
FILE REFERENCE: 58183W003  
CURRENT APPLICATION NUMBER: US/10/732,563  
CURRENT FILING DATE: 2003-12-10  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 6  
LENGTH: 904  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-732-563-6

Query Match 89.8%; Score 114; DB 16; Length 904;  
Best Local Similarity 96.0%; Pred. No. 4,1e-09;  
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FWLEERDFEAGVFELEAIYNSIKRS 25  
DB 786 FCLERDFEAGVFELEAIYNSIKRS 810

RESULT 6  
US-10-732-796A-6  
Sequence 6, Application US/10732796A  
Publication No. US20040197865A1  
GENERAL INFORMATION:  
APPLICANT: Gupta, Shalley K.  
APPLICANT: Ghosh, Tarun K.  
APPLICANT: Pink, Jason R.  
TITLE OF INVENTION: Gene Expression Systems and Recombinant Cell Lines  
FILE REFERENCE: 58182US004  
CURRENT APPLICATION NUMBER: US/10/732,796A  
CURRENT FILING DATE: 2003-12-10  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 6  
LENGTH: 904  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-732-796A-6

Query Match 89.8%; Score 114; DB 17; Length 904;  
 Best Local Similarity 96.0%; Pred. No. 4,1e-09;  
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FWLEERDFEAGVFELEAIYNSIKRS 25  
 DB 786 FWLEERDFEAGVFELEAIYNSIKRS 810

RESULT 7  
 US-10-265-072-4  
 ; Sequence 4, Application US/10265072  
 ; Publication No. US2003016001A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lipford, Grayson  
 ; TITLE OF INVENTION: TOLL-LIKE RECEPTOR 3 SIGNALING AGONISTS AND ANTAGONISTS  
 ; FILE REFERENCE: C01041.70031.US  
 ; CURRENT APPLICATION NUMBER: US/10/265,072  
 ; CURRENT FILING DATE: 2002-10-05  
 ; NUMBER OF SEQ ID NOS: 117  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 905  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-265-072-4

Query Match 79.5%; Score 101; DB 14; Length 905;  
 Best Local Similarity 88.0%; Pred. No. 4,8e-07;  
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FWLEERDFEAGVFELEAIYNSIKRS 25  
 DB 787 FWLEERDFEAGVFELEAIYNSIKRS 811

RESULT 8  
 US-10-407-952-22  
 ; Sequence 22, Application US/10407952  
 ; Publication No. US20030232074A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lipford, Grayson  
 ; APPLICANT: Bauer, Stefan  
 ; TITLE OF INVENTION: Immunostimulatory G,U-Containing Oligoribonucleotides  
 ; FILE REFERENCE: C01041.70037.US  
 ; CURRENT APPLICATION NUMBER: US/10/407,952  
 ; CURRENT FILING DATE: 2003-04-04  
 ; PRIOR APPLICATION NUMBER: US 60/421,966  
 ; PRIOR FILING DATE: 2002-10-29  
 ; PRIOR APPLICATION NUMBER: US 60/370,515  
 ; PRIOR FILING DATE: 2002-04-04  
 ; NUMBER OF SEQ ID NOS: 39  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 22  
 ; LENGTH: 905  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-407-952-22

Query Match 79.5%; Score 101; DB 14; Length 905;  
 Best Local Similarity 88.0%; Pred. No. 4,8e-07;  
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FWLEERDFEAGVFELEAIYNSIKRS 25  
 DB 787 FWLEERDFEAGVFELEAIYNSIKRS 811

RESULT 9  
 US-10-282-122A-57969  
 ; Sequence 57969, Application US/10282122A  
 ; Publication No. US20040029129A1

GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyckind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 57969  
 ; LENGTH: 367  
 ; TYPE: PRT  
 ; ORGANISM: Enterococcus faecium  
 US-10-282-122A-57969

Query Match 46.5%; Score 59; DB 15; Length 367;  
 Best Local Similarity 37.5%; Pred. No. 0.84;  
 Matches 9; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 2 WLEERDFEAGVFELEAIYNSIKRS 25  
 DB 123 WLEERDFEAGVFELEAIYNSIKRS 146

RESULT 10  
 US-10-424-599-149407  
 ; Sequence 149407, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J  
 ; APPLICANT: Kovalic, David K  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-26  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 149407  
 ; LENGTH: 117  
 ; TYPE: PRT

```

; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(117)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105938C.1.pep
US-10-424-599-149407

Query Match          39.4%; Score 50; DB 15; Length 117;
Best Local Similarity 48.1%; Pred. No. 6.3;
Matches 13; Conservative 2; Mismatches 6; Indels 6; Gaps 1;

QY 2 WLE-----RDFAGVFELAIYNSI 22
   |||:|||||:|||||:|||||:
DB 62 WLEDNQEADKEDFEAKQKEVEKIVNPI 88

RESULT 11
US-10-369-493-5415
; Sequence 5415, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianteng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5415
; LENGTH: 925
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5415

Query Match          38.6%; Score 49; DB 14; Length 925;
Best Local Similarity 32.3%; Pred. No. 94;
Matches 10; Conservative 7; Mismatches 6; Indels 8; Gaps 1;

QY 2 WLE-----RDFAGVFELAIYNSIKR 24
   |||:|||||:|||||:|||||:
DB 738 WLEDVTKDPTPKDFTDNLBLKVNVRSVKX 768

RESULT 12
US-10-282-122A-72448
; Sequence 72448, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078

```

```

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72448
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72448

Query Match          37.8%; Score 48; DB 15; Length 651;
Best Local Similarity 28.6%; Pred. No. 91;
Matches 12; Conservative 7; Mismatches 5; Indels 18; Gaps 1;

QY 2 WLEERDFEAGVPEL-----EAIYNSIKRS 25
   |||:|||||:|||||:|||||:
DB 547 WLOEKIESAVYEMCDMLLTNEVSIKKYRAELAIMSKRS 588

RESULT 13
US-10-406-686A-23
; Sequence 23, Application US/10406686A
; Publication No. US20040033586A1
; GENERAL INFORMATION:
; APPLICANT: CROOKE, HELEN RACHEL
; APPLICANT: SHEA, JACQUELINE ELIZABETH
; APPLICANT: FELDMAN, ROBERT GRAHAM
; APPLICANT: COUTERBROZE, SYLVAIN GABRIEL
; APPLICANT: LEGROS, FRANCOIS-XAVIER
; TITLE OF INVENTION: ATTENUATED GRAM NEGATIVE BACTERIA
; FILE REFERENCE: 454313-3171.1
; CURRENT APPLICATION NUMBER: US/10/406,686A
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/370,282
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-406-686A-23

Query Match          37.8%; Score 48; DB 15; Length 713;
Best Local Similarity 36.0%; Pred. No. 1e+02;
Matches 9; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 FWLEERDFEAGVPELAIYNSIKRS 25
   |||:|||||:|||||:|||||:
DB 4 FFVIRKDSGRTGFEIORTIIMAIKKA 28

RESULT 14
US-09-815-242-13637
; Sequence 13637, Application US/09815242

```

```
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13637
LENGTH: 666
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13637
```

```
Query Match 37.4% Score 47.5; DB 9; Length 666;
Best Local Similarity 50.0%; Pred. No. 1,1e+02;
Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 1;
```

```
QY 6 RDF--EAGVFELEAIIVNSIKR 24
Db 468 KDFPNLEYGDFEIEQILINIVKR 469

RESULT 15
US-09-815-242-13513
Sequence 13513, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
```

```
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13513
LENGTH: 676
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13513
```

```
Query Match 37.4% Score 47.5; DB 9; Length 676;
Best Local Similarity 50.0%; Pred. No. 1,1e+02;
Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 1;
```

Search completed: November 10, 2004, 16:35:47  
Job time : 25.1127 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:36:51 ; Search time 5.05401 Seconds  
(without alignments)  
475.942 Million cell updates/sec

Title: US-10-092-750-11  
Perfect score: 127  
Sequence: 1 FMLEERDFEAGFELFAVNSIKRS 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pirl1.\*  
2: pirl2.\*  
3: pirl3.\*  
4: pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	40.2	180	2 H90310	hypothetical prote
2	50	39.4	273	2 A72378	conserved hypothet
3	50	39.4	643	2 P72284	pled-related prote
4	49	38.6	925	2 G88175	protein T247.2 [1
5	48	37.8	317	2 T35981	probable peptide c
6	48	37.8	658	2 S38890	dnak-type molecula
7	48	37.8	707	1 A64047	ribonucleoside-tri
8	47.5	37.4	273	2 A99963	conserved hypothet
9	47.5	37.4	676	2 D95249	transcription regu
10	47.5	37.4	676	2 A98114	conserved hypothet
11	47	37.0	408	2 H71672	amg protein (amg)
12	47	37.0	102	2 AF2512	hypothetical prote
13	46	36.2	506	2 A81191	hypothetical prote
14	46	36.2	649	2 H97891	DNA mismatch repai
15	46	36.2	649	2 A95020	DNA mismatch repai
16	46	36.2	649	2 A93589	mismatch repair pr
17	46	36.2	933	2 F81989	hypothetical prote
18	46	36.2	933	2 F91195	probable integrase
19	45.5	35.8	393	2 G86042	probable integrase
20	45.5	35.8	393	2 T02698	Nicotiana tabacum
21	45.5	35.8	465	2 T35763	probable aminopept
22	45.5	35.8	835	2 T10195	hypothetical prote
23	45	35.4	141	2 S31768	cystathionine beta
24	45	35.4	178	2 S75999	hypothetical prote
25	45	35.4	228	2 T45396	Fixa [imported] -
26	45	35.4	266	2 T45396	hypothetical prote
27	45	35.4	314	2 T47882	L-lactate dehydrog
28	45	35.4	320	2 T47882	hypothetical prote
29	45	35.4	379	2 C84253	hypothetical prote

30	45	35.4	591	2 UC7657	mannan endo-1,4-be
31	45	35.4	652	2 A53163	dnak-type molecula
32	45	35.4	1189	2 UC6118	SH2-containing pro
33	45	35.4	1245	2 D86260	protein T12C24.22
34	44.5	35.0	537	2 A81869	type I site-specif
35	44.5	35.0	942	2 S75598	poly(A) polymerase
36	44	34.6	167	2 T35368	hypothetical prote
37	44	34.6	194	2 T33345	hypothetical prote
38	44	34.6	252	2 G70924	hypothetical prote
39	44	34.6	290	2 S21877	dnak-type molecula
40	44	34.6	343	2 C64555	3-dehydroquinatase
41	44	34.6	343	2 E71952	3-dehydroquinatase
42	44	34.6	386	2 T16099	hexuronate transpo
43	44	34.6	438	2 H87371	glycine-tRNA ligase
44	44	34.6	449	1 S73808	hypothetical prote
45	44	34.6	481	2 B95178	hypothetical prote

## ALIGNMENTS

## RESULT 1

H90310 hypothetical protein SS01521 [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C/Accession: H90310

R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-Y

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.

submitted to Genbank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: H90310

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-180 <CUR>

A/Cross-references: UNIPROT:Q97Y28; GB:AE006641; NID:g13814753; PIDN:AAK1743.1; GSPDB:GI

A/Gene: SS01521

Query Match 40.2%; Score 51; DB 2; Length 180;  
Best Local Similarity 48.0%; Pred. No. 2.2;  
Matches 12; Conservativity 3; Mismatches 8; Indels 2; Gaps 1;

QY 2 WLEF--RDFEAGFELFAVNSIKR 24

Db 156 WVEIEDFNANVTKLERLTNIKR 180

## RESULT 2

A72378 conserved hypothetical protein - Thermotoga maritima (strain MSB8)

C/Species: Thermotoga maritima

C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C/Accession: A72378

R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.R.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M. Nature 399, 323-329, 1999

A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A/Reference number: A72200; WUID:99287316; PMID:10360571

A/Accession: A72378

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-273 <ARN>

A/Cross-references: UNIPROT:Q9WYB8; GB:AE001721; GB:AE000512; NID:g4980922; PIDN:AAD3550

A/Experimental source: strain MSB8

C/Genetics:

A/Gene: TM0417

Query Match 39.4%; Score 50; DB 2; Length 273;  
Best Local Similarity 56.2%; Pred. No. 5.1;





A:Molecule type: DNA  
A:Residues: 1-707 <TIGR>  
A:Cross-references: UNIPROT:P43752; GB:U32693; GB:I42023; NID:G157021; PIDN:AAK21751.1;  
C:Comment: This enzyme must be activated by the anaerobic ribonucleotide reductase active  
peptide cleavage.  
C:Function:  
A:Description: catalyzes the reduction by reduced thioredoxin of a ribonucleoside tripho  
A:Pathway: deoxyribonucleotide biosynthesis  
A:Note: this enzyme is different both from ribonucleoside-triphosphate reductase (EC 1.1  
nds iron and has a tyrosyl radical  
C:Superfamily: Escherichia coli oxygen-sensitive ribonucleoside-triphosphate reductase;  
Ide-triphosphate reductase middle homology; rubredoxin homology  
C:Keywords: deoxyribonucleotide biosynthesis; iron; metalloprotein; oxidoreductase  
F:99-282/Domain: oxygen-sensitive ribonucleoside-triphosphate reductase middle homology  
F:308-707/Domain: oxygen-sensitive ribonucleoside-triphosphate reductase carboxyl-termin  
F:642-672/Domain: rubredoxin homology #status atypical <RUB>  
F:645-648,663,666/Binding site: iron (Cys) #status predicted  
F:682/Active site: Gly (stable glycol radical) #status predicted

Query Match 37.4%; Score 48; DB 1; Length 707;  
Best Local Similarity 36.0%; Pred. No. 31;  
Matches 9; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

QY 1 FWLEERDEAGVFELEAIYNSIKRS 25  
DB 4 FGVIRKDSGRAFEIQRITNAIKKA 28

RESULT 8  
A99963  
conserved hypothetical protein SA2362 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: A99963  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Uti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratake, K.  
L:ncet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A99788; MUID:21511952; PMID:11418146  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-273 <KUR>  
A:Cross-references: UNIPROT:Q99R62; GB:BA000018; PID:G13702526; PIDN:BAB43667.1; GSPDB:C  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: SA2362

Query Match 37.4%; Score 47.5; DB 2; Length 273;  
Best Local Similarity 45.8%; Pred. No. 12;  
Matches 11; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 2 WLEBRDFEAGVFELEAIYNSIKRS 25  
DB 239 WLEKGEYVD-FELMAEINPRLS 261

RESULT 9  
D95249  
transcription regulator, BglG family [imported] - Streptococcus pneumoniae (strain TIGR4  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: D95249  
R:Petzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
neon, T.; Hikey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21537209; PMID:11463916  
A:Accession: D95249  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-676 <KUR>  
A:Cross-references: UNIPROT:Q97NCO; GB:AE005672; PIDN:AAK76189.1; PID:G14973643; GSPDB:G  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP2131

Query Match 37.4%; Score 47.5; DB 2; Length 676;  
Best Local Similarity 50.0%; Pred. No. 35;  
Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 6 RDP---EAGVFELEAIYNSIKR 24  
DB 478 KDFENLRYGDFEIQIINIVKR 499

RESULT 10  
A98114  
conserved hypothetical protein spr1940 [imported] - Streptococcus pneumoniae (strain Re)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A98114  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E  
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain Re.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A98114  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-676 <KUR>  
A:Cross-references: UNIPROT:Q9BN39; GB:AE007317; PIDN:AAU00742.1; PID:G15459638; GSPDB:G  
A:Genetics:  
A:Gene: spr1940

Query Match 37.4%; Score 47.5; DB 2; Length 676;  
Best Local Similarity 50.0%; Pred. No. 35;  
Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 6 RDP---EAGVFELEAIYNSIKR 24  
DB 478 KDFENLRYGDFEIQIINIVKR 499

RESULT 11  
E87714  
hypothetical protein CC3751 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: E87714  
R:Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Duxkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, U.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: E87714  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-350 <STO>  
A:Cross-references: UNIPROT:Q9A215; GB:AE005673; NID:G13425525; PIDN:AAK25713.1; GSPDB:C  
C:Genetics:  
A:Gene: CC3751

Query Match 37.0%; Score 47; DB 2; Length 350;  
Best Local Similarity 47.6%; Pred. No. 20;  
Matches 10; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

QY 1 FWLEERDF--EAGVFELEAIY 19  
DB 290 FWKQEREFELROARAWNLEAIL 310

## RESULT 12

H71672  
amp9 protein (ampG2) RP668 - Rickettsia prowazekii  
C/Species: Rickettsia prowazekii  
C/Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 09-Jul-2004  
C/Accession: H71672  
R/Author: S.G.E. Zomorodipour, A.; Andersson, U.O.; Sichertitz-Ponten, T.; Alsmark, U.  
Nature 396, 133-140, 1998  
A/Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A/Reference number: A71630; MUID:19039499; PMID:9823893  
A/Accession: H71672  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-408 <AND>  
A/Cross-references: UNIPROT:Q9ZC01; GB:AJ235272; GB:AJ235269; NID:G3861033; PIDN:CAAL510  
A/Experimental source: strain Madrid E  
C/Genetics:  
A/Genome: ampG2; RP668

Query Match 37.0%; Score 47; DB 2; Length 408;  
Best Local Similarity 45.5%; Pred. No. 24;  
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

OY 1 FWLEEDFPAAGVFELEAIYNSIKR 22  
DB 37 FWLEEDFPAAGVFELEAIYNSIKR 58

## RESULT 13

AF2512  
hypothetical protein alr7278 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a1  
C/Species: Nostoc sp. PCC 7120  
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C/Accession: AF2512  
R/Author: T. Nakamura, Y.; Molk, C.P.; Kuritz, T.; Saegmoe, S.; Marande, A.; Iriguchi, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 6, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
A/Reference number: AB1807; MUID:21595285; PMID:11759840  
A/Accession: AF2512  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-102 <KUR>  
A/Cross-references: UNIPROT:Q9YKJ6; GB:BA000020; PIDN:BA078362.1; PID:gl7135816; GSPDB:G  
A/Experimental source: strain PCC 7120  
C/Genetics:  
A/Genome: alr7278

Query Match 36.2%; Score 46; DB 2; Length 102;  
Best Local Similarity 45.8%; Pred. No. 6.9;  
Matches 11; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

OY 1 FWLEEDFPAAGVFELEAIYNSIKR 24  
DB 12 FWLEEDFPAAGVFELEAIYNSIKR 33

## RESULT 14

AB1191  
hypothetical protein NMB0506 [imported] - Neisseria meningitidis (strain MC58 serogroup C/Species: Neisseria meningitidis  
C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C/Accession: AB1191  
R/Author: H. Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; et al.; Qin, H.; Vamathevan, J.; Gail, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappunoli, R.; Vabnick, A.; et al.  
A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A/Reference number: AB1000; MUID:20175755; PMID:10710307  
A/Accession: AB1191  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-506 <LET>  
A/Cross-references: UNIPROT:Q9K0S2; GB:AE002407; GB:AE002098; NID:G7225726; PIDN:AAAF4093  
A/Experimental source: serogroup B, strain MC58  
C/Genetics:  
A/Genome: NMB0506

Query Match 36.2%; Score 46; DB 2; Length 506;  
Best Local Similarity 38.1%; Pred. No. 43;  
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 4 EERDFPAAGVFELEAIYNSIKR 24  
DB 307 QELHRAIIEIQAITNQR 327

## RESULT 15

H97891  
DNA mismatch repair protein [imported] - Streptococcus pneumoniae (strain R6)  
C/Species: Streptococcus pneumoniae  
C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C/Accession: H97891  
R/Author: J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; Eick, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaaskunas, S.R.; et al.  
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A/Reference number: A97872; MUID:21429245; PMID:11544234  
A/Accession: H97891  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-649 <KUR>  
A/Cross-references: UNIPROT:P14160; GB:AE007317; PIDN:AAK98964.1; PID:gl5457702; GSPDB:G  
A/Experimental source: strain R6  
C/Genetics:  
A/Genome: hexB

Query Match 36.2%; Score 46; DB 2; Length 649;  
Best Local Similarity 26.2%; Pred. No. 57;  
Matches 11; Conservative 7; Mismatches 6; Indels 18; Gaps 1;

OY 2 WLEEDFPAAGVFELEAIYNSIKR 25  
DB 545 WMAEBIEISGIYEMCDMLLTREVSIRKXRAELAIMSCRS 586

Search completed: November 10, 2004, 12:29:13  
Job time : 6.05401 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:27:01 ; Search time 27.3534 Seconds  
(with about 31 elements)

Title: US-10-092-750-11

Sequence: 1 FWLEERDFEAGVFLELAI VNSIKRS 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

```
Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	114	89.8	870	2	06PCD4	06pcd4 homo sapien
2	114	89.8	870	2	AAH5372	AAH5372 homo sapi
3	114	89.8	904	1	TLR3_HUMAN	TLR3_HUMAN
4	105	82.7	154	2	OBMIK4	OBMIK4 bos tauris
5	101	79.5	905	1	TLR3_MOUSE	TLR3_MOUSE
6	101	79.5	905	1	Q7TN18	Q7tn18 mus musculu
7	101	79.5	905	2	BAC3J082	Bac3j082 mus muscu
8	55	43.3	250	2	06FI32	06fi32 mesoplasma
9	53	41.7	287	2	06VFR5	06vf5 anopheles g
10	53	41.7	287	2	AAH1784	AAH1784 anopheles
11	53	41.7	287	2	AAH1785	AAH1785 anopheles
12	53	41.7	287	2	AAH1786	AAH1786 anopheles
13	53	41.7	287	2	AAH1787	AAH1787 anopheles
14	53	41.7	287	2	AAH1788	AAH1788 anopheles
15	53	41.7	576	2	08WRE3	08wr3 anopheles g
16	53	41.7	817	2	Q7OIN8	Q7oin8 anopheles g
17	53	41.7	920	2	OB1ID3	OB1id3 bacillus ce
18	52	40.9	900	2	Q5VPH1	Q5vph1 drosophila
19	52	40.9	927	2	Q6HFE5	Q6hfe5 bacillus th
20	51	40.2	180	2	Q57Y28	Q57y28 sulfolobus
21	51	40.2	180	2	Q6CA08	Q6ca08 yarrowia li
22	50	39.4	273	2	Q5WYF8	Q5wyf8 thermotoga
23	50	39.4	643	3	Q9XOR6	Q9xor6 thermotoga
24	50	39.4	927	2	Q732W2	Q732w2 bacillus ce
25	50	39.4	927	2	OB1A28	OB1a28 bacillus ce
26	50	39.4	927	2	OB1W62	OB1w62 bacillus ce
27	50	39.4	927	2	AAH42703	AAH42703 bacillus
28	50	39.4	927	2	AAH33009	AAH33009 bacillus
29	49	38.6	172	2	Q6GV21	Q6gv21 bos tauris
30	49	38.6	392	2	Q72GR0	Q72gr0 thermus t
31	49	38.6	392	2	AAH82190	AAH82190 thermus t

32	49	38.6	43.9	1	SYN_CLOTE
33	49	38.6	59.4	2	OBUFTO
34	49	38.6	70.4	3	OBUFTO
35	49	38.6	79.3	2	O704V6
36	49	38.6	79.3	2	O706D2
37	49	38.6	79.3	2	CAP02015
38	49	38.6	79.3	2	CAP02197
39	49	38.6	92.5	2	C22758
40	49	38.6	99.1	2	Q6RSN8
41	49	38.6	99.1	2	AA537674
42	48.5	38.2	21.3	2	O8XND8
43	48	37.8	21.5	2	O7UCN4
44	48	37.8	25.1	2	O6D426
45	48	37.8	31.7	2	O92533
46	48	37.8	36.5	2	Q97221
47	48	37.8	36.5	2	Q97221
48	48	37.8	36.5	2	Q97221
49	48	37.8	36.5	2	Q97221
50	48	37.8	36.5	2	Q97221
51	48	37.8	36.5	2	Q97221
52	48	37.8	36.5	2	Q97221
53	48	37.8	36.5	2	Q97221
54	48	37.8	36.5	2	Q97221
55	48	37.8	36.5	2	Q97221
56	48	37.8	36.5	2	Q97221
57	48	37.8	36.5	2	Q97221
58	48	37.8	36.5	2	Q97221
59	48	37.8	36.5	2	Q97221
60	48	37.8	36.5	2	Q97221
61	48	37.8	36.5	2	Q97221
62	48	37.8	36.5	2	Q97221
63	48	37.8	36.5	2	Q97221
64	48	37.8	36.5	2	Q97221
65	48	37.8	36.5	2	Q97221
66	48	37.8	36.5	2	Q97221
67	48	37.8	36.5	2	Q97221
68	48	37.8	36.5	2	Q97221
69	48	37.8	36.5	2	Q97221
70	48	37.8	36.5	2	Q97221
71	48	37.8	36.5	2	Q97221
72	48	37.8	36.5	2	Q97221
73	48	37.8	36.5	2	Q97221
74	48	37.8	36.5	2	Q97221
75	48	37.8	36.5	2	Q97221
76	48	37.8	36.5	2	Q97221
77	48	37.8	36.5	2	Q97221
78	48	37.8	36.5	2	Q97221
79	48	37.8	36.5	2	Q97221
80	48	37.8	36.5	2	Q97221
81	48	37.8	36.5	2	Q97221
82	48	37.8	36.5	2	Q97221
83	48	37.8	36.5	2	Q97221
84	48	37.8	36.5	2	Q97221
85	48	37.8	36.5	2	Q97221
86	48	37.8	36.5	2	Q97221
87	48	37.8	36.5	2	Q97221
88	48	37.8	36.5	2	Q97221
89	48	37.8	36.5	2	Q97221
90	48	37.8	36.5	2	Q97221
91	48	37.8	36.5	2	Q97221
92	48	37.8	36.5	2	Q97221
93	48	37.8	36.5	2	Q97221
94	48	37.8	36.5	2	Q97221
95	48	37.8	36.5	2	Q97221
96	48	37.8	36.5	2	Q97221
97	48	37.8	36.5	2	Q97221
98	48	37.8	36.5	2	Q97221
99	48	37.8	36.5	2	Q97221
100	48	37.8	36.5	2	Q97221

## ALIGNMENTS

RESULT 1	ID	PRELIMINARY;	PRT; 870 AA.
OCPCD4	OCPCD4		
AC	OCPCD4		
DT	05-JUL-2004 (TEMBLrel, 27, Created)		
DT	05-JUL-2004 (TEMBLrel, 27, Last sequence update)		
DT	05-JUL-2004 (TEMBLrel, 27, Last annotation update)		
DE	TLR3 protein (Fragment).		
GN	Name=TLR3;		
OS	Homo sapiens (human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Placenta;		
RY	MEDLINE=22388257; PubMed=12477932;		
RA	Strauberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Dickenson L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Sefton M., Soares M.B., Bonaldo M.F., Casavant T.J., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinzi P., Prange C.,		
RA	Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.O., Males J.A., Gunnarath P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,		
RA	Kryzhanovskii M.I., Szalska U., Smalins D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Placenta;		
RA	Strauberg R.;		
RL	Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC059372; AAH59372.1; -		
DR	InterPro; IPR001611; LRR.		
DR	InterPro; IPR000483; LRR Cyterm.		
DR	InterPro; IPR003885; LRR Cyct.		
DR	InterPro; IPR003594; LRR_CYP.		
DR	InterPro; IPR000157; TIR.		
DR	Pfam; PF00560; LRR_19.		
DR	Pfam; PF01463; LRRCT_1.		
DR	Pfam; PF01582; TIR_1.		
DR	PRINTS; PR00019; LEURICRAFT.		
DR	SMART; SMO00982; LRRCT_1.		
DR	SMART; SMO0363; LRR_SD22; 8.		

```

DR SMART; SM00369; LRR_TYP; 16.
DR SMART; SM00255; TIR; 1.
DR PROSITE; PSS0104; TIR; 1.
FT NON_TER 870
SQ SEQUENCE 870 AA; 99927 MW; 0404ED033885D9BC CRC64;

Query Match
Best Local Similarity 96.0%; Score 114; DB 2; Length 870;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FWLEERDFAGVFELEALVNSIKRS 25
DB 786 FCLERDFEAGVFELEALVNSIKRS 810

RESULT 2
ID AAH59372 PRELIMINARY; PRT; 870 AA.
AC AAH59372;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DE 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE TLR3 protein (Fragment).
GN TLR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Scheiner C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheeler C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max T., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Cantini P., Prange C.J.,
RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosa S.A., McKernan P.J., McKernan K.O., Malek J.A., Gunatane P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.W., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzyzanski M.I., Skalski U., Smallus D.B., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC059372; AAH59372.1; -
FT NON_TER 870
SQ SEQUENCE 870 AA; 99927 MW; 0404ED033885D9BC CRC64;

Query Match
Best Local Similarity 96.0%; Score 114; DB 2; Length 870;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FWLEERDFAGVFELEALVNSIKRS 25
DB 786 FCLERDFEAGVFELEALVNSIKRS 810

RESULT 3
TLR3_HUMAN STANDARD; PRT; 904 AA.
ID TLR3_HUMAN

```

---

```

AC O15455;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Toll-like receptor 3 precursor.
GN Name=TLR3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98118556; PubMed=9435236;
RA Rock F.L., Hardiman G., Timans J.C., Kastelein R.A., Bazan J.F.;
RT "A family of human receptors structurally related to Drosophila
RT Toll."
RL Proc. Natl. Acad. Sci. U.S.A. 95:598-593 (1998).
RN [2]
RP SEQUENCE OF 24-38.
RA Zhang Z., Henzel W.;
RT "Signal peptide prediction based on analysis of experimentally
RT verified cleavage sites."
RL Submitted (JUN-2004) to Swiss-Prot.
CC -!- FUNCTION: Participates in the innate immune response to microbial
CC agents. May be involved in the recognition of ds-RNA. Acts via
CC MyD88 and TRAF6, leading to NF-kappa-B activation, cytokine
CC secretion and the inflammatory response (By similarity).
CC -!- SUBUNIT: Binds MyD88 via their respective TIR domains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed at high level in placenta and
CC pancreas. Also detected in CD13c+ immature dendritic cells. Only
CC expressed in dendritic cells and not in other leukocytes,
CC including monocyte precursors.
CC -!- SIMILARITY: Belongs to the Toll-like receptor family.
CC -!- SIMILARITY: Contains 22 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 TIR domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U88879; AAC34134.1; -
DR HSPB; P07359; 1M0Z; TLR3.
DR Genew; HGNC:11849; TLR3.
DR GO; GO:0005887; C:Integral to plasma membrane; TAS.
DR GO; GO:0016020; C:membrane; NAS.
DR GO; GO:0003725; F:double-stranded RNA binding; NAS.
DR GO; GO:0004688; F:transmembrane receptor activity; NAS.
DR GO; GO:0007250; F:activation of NF-kappaB-inducing kinase; NAS.
DR GO; GO:0042742; P:defense response to bacteria; TAS.
DR GO; GO:0006972; P:detection of virus; NAS.
DR GO; GO:0006872; P:hyperosmotic response; NAS.
DR GO; GO:0045671; P:negative regulation of osteoclast different. . .; NAS.
DR GO; GO:0045359; P:positive regulation of interferon-beta bios. . .; IMP.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_TYP.
DR InterPro; IPR000157; TIR.
DR Pfam; PF00560; LRR; 18.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01582; TIR; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 2.

```

DR SMART; SM00255; TIR; 1.  
 DR PROSITE; PS50104; TIR; 1.  
 KW Direct protein sequencing; Glycoprotein; Immune response;  
 KW Inflammatory response; Leucine-rich repeat; Receptor; Repeat; Signal;  
 KW Transmembrane.

FT SIGNAL 1 23  
 FT CHAIN 24 904  
 FT DOMAIN 24 704  
 FT TRANSMEM 705 725  
 FT DOMAIN 726 904  
 FT REPEAT 50 73  
 FT REPEAT 75 97  
 FT REPEAT 98 121  
 FT REPEAT 123 145  
 FT REPEAT 147 169  
 FT REPEAT 170 193  
 FT REPEAT 196 219  
 FT REPEAT 247 270  
 FT REPEAT 273 295  
 FT REPEAT 298 320  
 FT REPEAT 354 377  
 FT REPEAT 379 403  
 FT REPEAT 406 429  
 FT REPEAT 431 454  
 FT REPEAT 455 478  
 FT REPEAT 480 504  
 FT REPEAT 505 528  
 FT REPEAT 530 552  
 FT REPEAT 561 584  
 FT REPEAT 586 608  
 FT REPEAT 610 632  
 FT REPEAT 634 659  
 FT DOMAIN 754 896  
 FT CAROCHD 52 52  
 FT CAROCHD 57 57  
 FT CAROCHD 70 70  
 FT CAROCHD 124 124  
 FT CAROCHD 196 196  
 FT CAROCHD 247 247  
 FT CAROCHD 252 252  
 FT CAROCHD 265 265  
 FT CAROCHD 275 275  
 FT CAROCHD 291 291  
 FT CAROCHD 398 398  
 FT CAROCHD 413 413  
 FT CAROCHD 507 507  
 FT CAROCHD 636 636  
 FT CAROCHD 662 662  
 SQ SEQUENCE 904 AA; 103828 MW; 034E05E6CA7A4D2F7 CRC64;

Query Match 89.8%; Score 114; DB 1; Length 904;  
 Best Local Similarity 96.0%; Pred. No. 1.9e-08;  
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 FMLERDFEAGVFEEALVNSIKRS 25  
 DB 786 FMLERDFEAGVFEEALVNSIKRS 810

RESULT 4  
 OSMTK4 PRELIMINARY; PRT; 154 AA.  
 AC OSMTK4  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE TOLL-like receptor 3 (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Werling D., Benedek E., Dubey C., Jung T.,  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY124007; AAM91032.1;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:004888; F:transmembrane receptor activity; IEA.  
 DR InterPro; IPR004075; IIL\_receptor1.  
 DR InterPro; IPR00157; TIR.  
 DR Pfam; PF01582; TIR; 1.  
 DR PRINTS; PR01537; INTERKNIRF.  
 DR SMART; SM00255; TIR; 1.  
 DR PROSITE; PS50104; TIR; 1.  
 KW Receptor.  
 FT NON-TER 1 1  
 SQ SEQUENCE 154 AA; 18326 MW; E3BAE3871A3CFA7 CRC64;

Query Match 82.7%; Score 105; DB 2; Length 154;  
 Best Local Similarity 88.0%; Pred. No. 6.5e-08;  
 Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CY 1 FMLERDFEAGVFEEALVNSIKRS 25  
 DB 36 FMLERDFEAGVFEEALVNSIKRS 60

RESULT 5  
 TIR3 MOUSE STANDARD; PRT; 905 AA.  
 ID TIR3\_MOUSE  
 AC Q99MB1; Q912M4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE TOLL-like receptor 3 precursor.  
 GN Name=Tlr3;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv;  
 RA MEDLINE=21519099; PubMed=11607032; DOI=10.1038/35099560;  
 RA Alexopoulos L., Holt A.C., Medzhitov R., Flavell R.A.,  
 RT "Recognition of double-stranded RNA and activation of NF-kappa B  
 by Toll-like receptor 3.";  
 RL Nature 413:732-738(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c x NIH; TISSUE=Macrophage;  
 RA Applequist S.E., Ujungren H.G.,  
 RT "Molecular cloning of mouse Toll-like receptor 3 cDNA."  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RA MEDLINE=22354683; PubMed=1246681; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuta M., Kasukawa T., Adachi U., Bono H., Kondo S.,  
 RA Nikaio I., Osato N., Saito R., Suzuki H., Yamada I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schirni L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Choitha C., Corbett L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.W.,  
 RA Koragaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pereira G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor W.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilmung L.G., Wyszynski-Boris A., Yanagisawa M., Yang T., Yang L.,  
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carinci P., Hayatsu N.,  
 RA Hirose K., Kishikawa T., Komori H., Nakamura M., Sakazume N., Sato K.,  
 RA Shitaki T., Waki K., Kawai U., Aizawa K., Aikawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.,  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573(2002).  
 CC -1- FUNCTION: Participates in the innate immune response to microbial  
 CC agents. May be involved in the recognition of ds-RNA. Acts via  
 CC MyD88 and TRAF6, leading to NF-kappa-B activation, cytokine  
 CC secretion and the inflammatory response (By similarity).  
 CC -1- SUBUNIT: Binds MyD88 via their respective TIR domains (By  
 CC similarity).  
 CC -1- SIMILARITY: Contains 22 leucine-rich (LRR) repeats.  
 CC -1- SIMILARITY: Contains 1 TIR domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL, AF420279; AAL27007.1; -;  
 DR EMBL, AF355152; AAK26117.1; -;  
 DR EMBL, AK083977; BAC39082.1; -;  
 DR MGI, MGI:2156367; Tlr3.  
 DR GO, GO:0016020; C:membrane; ISS.  
 DR GO, GO:0003725; F:double-stranded RNA binding; ISS.  
 DR GO, GO:0004888; F:transmembrane receptor activity; ISS.  
 DR GO, GO:0007250; P:activation of NF-kappaB-inducing kinase; ISS.  
 DR GO, GO:0006952; P:defense response; IMP.  
 DR GO, GO:0006957; P:detection of virus; ISS.  
 DR GO, GO:0006972; P:hyperosmotic response; ISS.  
 DR GO, GO:0045671; P:negative regulation of osteoclast different. . .; ISS.  
 DR GO, GO:0045359; P:positive regulation of interferon-beta bios. . .; ISS.  
 DR InterPro, IPR001611; LRR.  
 DR InterPro, IPR000483; LRR\_Cterm.  
 DR InterPro, IPR000372; LRR\_Nterm.  
 DR InterPro, IPR003591; LRR\_Typ.  
 DR Pfam, PF00560; LRR; 18.  
 DR Pfam, PF01463; LRRCT; 1.  
 DR Pfam, PF01582; TIR; 1.  
 DR PRINTS, PR00019; LEURICRPT.  
 DR SMART, SM00082; LRRCT; 1.  
 DR SMART, SM00013; LRRNT; 1.  
 DR SMART, SM00369; LRR\_Typ; 4.  
 DR SMART, SM00255; TIR; 1.  
 DR PROSITE, PS50104; TIR; 1.  
 KW Glycoprotein; Immune response; Inflammatory response;  
 KW Leucine-rich repeat; Receptor; Signal; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 905  
 FT DOMAIN 26 705  
 FT TRANSMEM 706 726  
 FT DOMAIN 727 905  
 FT REPEAT 51 74  
 FT REPEAT 76 98  
 FT REPEAT 99 122  
 FT REPEAT 124 146  
 FT REPEAT 148 170  
 LRR 5.  
 LRR 4.  
 LRR 3.  
 LRR 2.  
 LRR 1.  
 Potential.  
 Extracellular (Potential).  
 Potential.  
 Cytoplasmic (Potential).

FT REPEAT 171 194  
 FT REPEAT 197 220  
 FT REPEAT 248 274  
 FT REPEAT 274 297  
 FT REPEAT 298 321  
 FT REPEAT 355 378  
 FT REPEAT 380 404  
 FT REPEAT 407 430  
 FT REPEAT 432 455  
 FT REPEAT 456 479  
 FT REPEAT 481 505  
 FT REPEAT 506 529  
 FT REPEAT 531 553  
 FT REPEAT 562 585  
 FT REPEAT 587 609  
 FT REPEAT 610 633  
 FT REPEAT 635 660  
 FT DOMAIN 755 897  
 FT CARBOHYD 53 53  
 FT CARBOHYD 58 58  
 FT CARBOHYD 71 71  
 FT CARBOHYD 125 125  
 FT CARBOHYD 197 197  
 FT CARBOHYD 248 248  
 FT CARBOHYD 253 253  
 FT CARBOHYD 276 276  
 FT CARBOHYD 292 292  
 FT CARBOHYD 399 399  
 FT CARBOHYD 414 414  
 FT CARBOHYD 425 425  
 FT CARBOHYD 508 508  
 FT CARBOHYD 653 653  
 FT CONFIDENCE 670 670  
 SQ SEQUENCE 905 AA; 103670 MW; 8E6ADE9818E1484 CRC64;  
 Query Match 79.5%; Score 101; DB 1; Length 905;  
 Best Local Similarity 88.0%; Pred. No. 1.9e-06;  
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 FMLERDFEAGVPELFAINSIKRS 25  
 Db 787 FMLERDFEAGVGLGFAIVNSIKRS 811  
 RESULT 6  
 Q7TN18 PRELIMINARY; PRT; 905 AA.  
 ID Q7TN18  
 AC Q7TN18  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Toll-like receptor 3.  
 GN Name-Tlr3;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RA Tamura H., Oikawa A., Kato H.,  
 RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AB116229; BAC81504.1; -;  
 DR GO, GO:0016020; C:membrane; IEA.  
 DR GO, GO:0004888; F:transmembrane receptor activity; IEA.  
 DR InterPro, IPR001611; LRR.  
 DR InterPro, IPR000483; LRR\_Cterm.  
 DR InterPro, IPR000372; LRR\_Nterm.  
 DR InterPro, IPR003591; LRR\_Typ.  
 DR InterPro, IPR000157; TIR.  
 DR Pfam, PF00560; LRR; 17.  
 DR Pfam, PF01463; LRRCT; 1.  
 DR Pfam, PF01582; TIR; 1.

```

DR SMART; SMO0082; LRRT; 1.
DR SMART; SMO0013; LRRT; 1.
DR SMART; SMO0369; LRRT; 4.
DR SMART; SMO0255; TIR; 1.
DR PROSITE; PS50104; TIR; 1.
KW Receptor.
SQ SEQUENCE 905 AA; 103174 MW; 478326F3A695D3D1 CRC64;

Query Match 79.5%; Score 101; DB 2; Length 905;
Best Local Similarity 88.0%; Pred. No. 1.9e-06;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 FWLEERDFEAGVLEIAVNSIKRS 25
Db 787 FCLERDFEAGVLEIAVNSIKRS 811

RESULT 7
BAC39082 PRELIMINARY; PRT; 905 AA.
AC BAC39082;
ID BAC39082;
DT 14-APR-2004 (TREMBlrel. 27, Created)
DT 14-APR-2004 (TREMBlrel. 27, Last sequence update)
DT 14-APR-2004 (TREMBlrel. 27, Last annotation update)
DE 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched
DE library, clone:DI30071D16 product:coll-1like receptor 3, full insert
DE sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=22354683; PubMed=14466851;
RA The FANTOM Consortium;
RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RL "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44 (1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuwa T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneida Y., Ishikawa T., Ozawa K., Tanaka T., Matsutani S., Kawai J.,

```

```

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RA Adachi U., Aizawa K., Akimura T., Aikawa T., Hasegawa W.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume M.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Kato F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kori H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Osato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK083977; BAC39082.1; -.
KW Receptor.
SQ SEQUENCE 905 AA; 103670 MW; 8EA6DBA9818E14B4 CRC64;

Query Match 79.5%; Score 101; DB 2; Length 905;
Best Local Similarity 88.0%; Pred. No. 1.9e-06;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 FWLEERDFEAGVLEIAVNSIKRS 25
Db 787 FCLERDFEAGVLEIAVNSIKRS 811

RESULT 8
O6F132 PRELIMINARY; PRT; 250 AA.
ID O6F132;
AC O6F132;
DT 01-OCT-2004 (TREMBlrel. 28, Created)
DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Transcriptional regulator.
DE ORFNames=MF1432;
GN Mesoplasma florum J1.
OS Mesoplasma florum J1.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmales;
OC Entomoplasmataceae; Mesoplasma.
OX NCBI_TaxID=265311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L1;
RA Bitren B.W., Strange-Thomann N., Hafez N., Decaprio D., Fisher S.,
RA Butler J., Elkins T., Kodira C.D., Major J., Wang S., Nicol R.,
RA Nusbam C.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=L1;
RA Knight T., Jr., Fournier G.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017263; AAT75791.1; -.
SQ SEQUENCE 250 AA; 29989 MW; 207DBFB12CDD1BFB CRC64;

Query Match 43.3%; Score 55; DB 2; Length 250;
Best Local Similarity 40.3%; Pred. No. 5.7;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Oy 3 LBEEDFEAGVLEIAVNSIKR 24
Db 143 LKEDYQAMIFNDEKLNKKK 164

RESULT 9
O6VFB5 PRELIMINARY; PRT; 287 AA.
ID O6VFB5;
AC O6VFB5;

```

```

DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE TOLL9 (Fragment).
GN Name=TOLL9.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J3-5; Yaounde, and 4ARR;
RA Morlais I.; Poncon N.; Simard F.; Cohuet A.; Fontenille D.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY341221; AAR13785.1; -
DR EMBL; AY341222; AAR13786.1; -
DR EMBL; AY341223; AAR13787.1; -
DR EMBL; AY341224; AAR13788.1; -
DR EMBL; AY341220; AAR13784.1; -
DR InterPro; IPR000157; TIR.
DR PROSITE; PSS0104; TIR; 1.
FT NON_TER 1
FT NON_TER 287
SQ SEQUENCE 287 AA; 32927 MW; A84166519456AA4E CRC64;

Query Match 41.7%; Score 53; DB 2; Length 287;
Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 LERDPFAGVFELIAIVNSIKRS 25
DB 238 LHERDFEVGYILENIISCMDS 260

RESULT 10
AAR13784 PRELIMINARY; PRT; 287 AA.
ID AAR13784
AC AAR13784;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE TOLL9 (Fragment).
GN TOLL9.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4ARR;
RA Morlais I.; Poncon N.; Simard F.; Cohuet A.; Fontenille D.;
RL "Intraspecific nucleotide variation in Anopheles gambiae: benefits and drawbacks.";
RT Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY341220; AAR13784.1; -
FT NON_TER 1
FT NON_TER 287
SQ SEQUENCE 287 AA; 32927 MW; A84166519456AA4E CRC64;

Query Match 41.7%; Score 53; DB 2; Length 287;
Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 LERDPFAGVFELIAIVNSIKRS 25
DB 238 LHERDFEVGYILENIISCMDS 260

RESULT 11
AAR13785 PRELIMINARY; PRT; 287 AA.
ID AAR13785
AC AAR13785;
DT 02-MAR-2004 (TREMBlrel. 27, Created)

```

```

DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE TOLL9 (Fragment).
GN TOLL9.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J3-5;
RA Morlais I.; Poncon N.; Simard F.; Cohuet A.; Fontenille D.;
RL "Intraspecific nucleotide variation in Anopheles gambiae: benefits and drawbacks.";
RT Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY341221; AAR13785.1; -
FT NON_TER 1
FT NON_TER 287
SQ SEQUENCE 287 AA; 32927 MW; A84166519456AA4E CRC64;

Query Match 41.7%; Score 53; DB 2; Length 287;
Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 LERDPFAGVFELIAIVNSIKRS 25
DB 238 LHERDFEVGYILENIISCMDS 260

RESULT 12
AAR13786 PRELIMINARY; PRT; 287 AA.
ID AAR13786
AC AAR13786;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE TOLL9 (Fragment).
GN TOLL9.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Yaounde;
RA Morlais I.; Poncon N.; Simard F.; Cohuet A.; Fontenille D.;
RL "Intraspecific nucleotide variation in Anopheles gambiae: benefits and drawbacks.";
RT Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY341222; AAR13786.1; -
FT NON_TER 1
FT NON_TER 287
SQ SEQUENCE 287 AA; 32927 MW; A84166519456AA4E CRC64;

Query Match 41.7%; Score 53; DB 2; Length 287;
Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 LERDPFAGVFELIAIVNSIKRS 25
DB 238 LHERDFEVGYILENIISCMDS 260

RESULT 13
AAR13787 PRELIMINARY; PRT; 287 AA.
ID AAR13787
AC AAR13787;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE TOLL9 (Fragment).
GN TOLL9.
OS Anopheles gambiae (African malaria mosquito).

```



```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicidae; Anopheles.
CX NCB1_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Yaounde;
RA Morlais I., Poncon N., Simard F., Cohuet A., Fontenille D.;
RT "Intraspecific nucleotide variation in Anopheles gambiae: benefits and
RT drawbacks";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY341223; AAR13787.1; -.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 287 AA; 32927 MW; A64166519456AA4E CRC64;

Cy 3 LEERDFEAGVFELFAIVNSIKRS 25
Db 238 LHERDFEVGYGILENTISCMDS 260

RESULT 14
AAR13788 PRELIMINARY; PRT; 287 AA.
AC AAR13788;
DT 02-MAR-2004 (TRENBLREL. 27, Created)
DT 02-MAR-2004 (TRENBLREL. 27, Last sequence update)
DT 02-MAR-2004 (TRENBLREL. 27, Last annotation update)
DE TOLL9 (Fragment).
GN TOLL9.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicidae; Anopheles.
CX NCB1_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4ARR;
RA Morlais I., Poncon N., Simard F., Cohuet A., Fontenille D.;
RT "Intraspecific nucleotide variation in Anopheles gambiae: benefits and
RT drawbacks";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY341224; AAR13788.1; -.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 287 AA; 32927 MW; A64166519456AA4E CRC64;

Query Match 41.7%; Score 53; DB 2; Length 287;
Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Cy 3 LEERDFEAGVFELFAIVNSIKRS 25
Db 238 LHERDFEVGYGILENTISCMDS 260

RESULT 15
Q8WRE3 PRELIMINARY; PRT; 576 AA.
AC Q8WRE3;
DT 01-MAR-2002 (TRENBLREL. 20, Created)
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
DE TOLL9.
GN Name=TOLL9;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicidae; Anopheles.
CX NCB1_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=22202213; PubMed=12213252;
RA Luna C., Wang X., Huang Y., Zhang J., Zheng L.;
RT "Characterization of four Toll related genes during development and
RT immune responses in Anopheles gambiae";
RL Insect Biochem. Mol. Biol. 32:1171-1179(2002).
DR EMBL; AF444782; AAL37903.1; -.
DR HSSP; Q15399; 1FYV.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:004888; F:transmembrane receptor activity; IEA.
DR InterPro; IPR004075; IL1_receptor1.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_TYP.
DR InterPro; IPR000157; TIR.
DR Pfam; PF00560; LRR_5.
DR Pfam; PF01582; TIR_1.
DR PRINTS; PRO1537; INTRUKN1RF.
DR PRINTS; PRO0019; LEURCHRT.
DR SMART; SM00369; LRR_TYP; 1.
DR SMART; SM00255; TIR_1.
DR PROSITE; PS50104; TIR_1.
SQ SEQUENCE 576 AA; 66800 MW; ECTE978E8AB3A737 CRC64;

Query Match 41.7%; Score 53; DB 2; Length 576;
Best Local Similarity 52.2%; Pred. No. 29;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Cy 3 LEERDFEAGVFELFAIVNSIKRS 25
Db 465 LHERDFEVGYGILENTISCMDS 487

```

Search completed: November 10, 2004, 12:27:04  
 Job time : 30.3534 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 ; Search time 11.9136 Seconds  
(without alignments)  
222.664 Million cell updates/sec

Title: US-10-092-750-12  
Perfect score: 1 MKMTDNTLGTETISWENKLAEGKLTLDITFVHHVLAAPH 40  
Sequence: 1 MKMTDNTLGTETISWENKLAEGKLTLDITFVHHVLAAPH 40

Scoring table: BLOSUM62  
Gap: 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Issued\_Patents\_AA\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backlist.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160	72.4	105	4	US-09-513-999C-5338
2	160	72.4	283	1	US-08-726-227-1
3	160	72.4	283	1	US-08-726-227-4
4	160	72.4	283	3	US-08-826-246-6
5	160	72.4	283	3	US-08-944-495-6
6	160	72.4	283	3	US-09-126-640-9
7	160	72.4	283	3	US-08-925-588-6
8	160	72.4	283	3	US-09-288-292A-9
9	160	72.4	283	4	US-09-372-044-6
10	160	72.4	283	4	US-08-825-486-6
11	160	72.4	283	4	US-08-826-248-6
12	112	50.7	292	1	US-08-726-227-3
13	53	24.0	383	4	US-09-489-039A-13424
14	53	24.0	822	4	US-09-538-092-941
15	53	24.0	822	4	US-09-538-092-941
16	52.5	23.8	254	4	US-09-540-235-2733
17	52.5	23.8	254	4	US-09-540-235-2733
18	51.5	23.3	286	3	US-09-513-999C-990
19	51.5	23.3	286	3	US-09-172-952-31
20	51.5	23.3	637	1	US-08-350-325A-10
21	51.5	23.3	637	5	PCT-US94-03856-10
22	51	23.1	132	4	US-09-583-110-4168
23	50.5	22.9	242	4	US-09-489-039A-10443
24	50.5	22.9	242	4	US-09-489-039A-10443
25	50	22.6	119	4	US-07-672-483-1
26	50	22.6	119	4	US-09-513-999C-5568
27	50	22.6	137	4	US-10-101-464A-664
			440	4	US-09-631-603-22

28	50	22.6	557	4	US-09-138-277C-3	Sequence 3, Appl.
29	50	22.6	707	4	US-09-252-991A-24334	Sequence 24334, A
30	50	22.6	823	4	US-09-886-319A-63	Sequence 63, Appl
31	49.5	22.4	380	3	US-09-181-338-17	Sequence 17, Appl
32	49.5	22.4	535	4	US-09-435-733-44	Sequence 44, Appl
33	49.5	22.4	535	4	US-09-930-218-1	Sequence 1, Appl
34	49.5	22.4	535	4	US-09-988-113-44	Sequence 44, Appl
35	49	22.2	275	4	US-09-489-039A-8651	Sequence 8651, Ap
36	49	22.2	564	4	US-09-107-532A-6236	Sequence 6236, Ap
37	49	22.2	564	4	US-09-107-532A-6237	Sequence 6237, Ap
38	49	22.2	992	3	US-08-890-865A-1	Sequence 1, Appl
39	48.5	21.9	192	4	US-09-543-681A-7846	Sequence 7846, Ap
40	48.5	21.9	420	3	US-09-255-358-6	Sequence 6, Appl
41	48.5	21.9	420	4	US-09-405-558-6	Sequence 6, Appl
42	48.5	21.9	486	1	US-07-672-483-2	Sequence 2, Appl
43	48.5	21.9	486	4	US-08-259-451-13	Sequence 13, Appl
44	48.5	21.9	599	4	US-09-328-352-7349	Sequence 7349, Ap
45	48	21.7	325	4	US-09-543-681A-7840	Sequence 7840, Ap

## ALIGNMENTS

```

RESULT 1
US-09-513-999C-5338
Sequence 5338, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclet, A. Y.
APPLICANT: Giordano, J. Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59 US2, REG
CURRENT APPLICATION NUMBER: US/09/513, 999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 5338
LENGTH: 105
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 14
OTHER INFORMATION: Xaa=Ala or Val
US-09-513-999C-5338

Query Match 72.4%; Score 160; DB 4; Length 105;
Best Local Similarity 96.8%; Pred. No. 1.3e-15;
Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 2 KMTDNTLGTETISWENKLAEGKLTLDITFV 32
Db 75 KMTDNTLGTETISWENKLAEGKLTLDITFV 105

RESULT 2
US-08-726-227-1
Sequence 1, Application US/08726227
Patent No. 5780235
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN VOLTAGE-DEPENDENT
TITLE OF INVENTION: ANION CHANNEL HACH
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto

```

STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/726,227  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0135 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-845-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ. ID NO. 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 283 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE: Consensus

US-08-726-227-1

Query Match 72.4%; Score 160; DB 1; Length 283;  
Best Local Similarity 96.8%; Pred. No. 4.5e-15;  
Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KWDNTLTGTEISWENKLAEGKLTLDITFV 32  
DB 74 KWDNTLTGTEISWENKLAEGKLTLDITFV 104

RESULT 3  
US-08-726-227-4  
Sequence 4, Application US/08726227  
Patent No. 5780235  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: NOVEL HUMAN VOLTAGE-DEPENDENT  
TITLE OF INVENTION: ANION CHANNEL HACH  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/726,227  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0135 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ. ID NO. 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 283 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1463028

US-08-726-227-4

Query Match 72.4%; Score 160; DB 1; Length 283;  
Best Local Similarity 96.8%; Pred. No. 4.5e-15;  
Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KWDNTLTGTEISWENKLAEGKLTLDITFV 32  
DB 74 KWDNTLTGTEISWENKLAEGKLTLDITFV 104

RESULT 4  
US-08-826-246-6  
Sequence 6, Application US/08826246  
Patent No. 6048709  
GENERAL INFORMATION:  
APPLICANT: Faib, Dean  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: THE TREATMENT AND DIAGNOSIS OF  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/826,246  
FILING DATE: 28-MAR-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/799,910  
FILING DATE: 13-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/011,787  
FILING DATE: 16-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Cortuzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-078-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)7909090  
TELEFAX: (212)8699741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ. ID NO. 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 283 amino acids  
TYPE: amino acid